

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2011 年 9 月 22 日(22.09.2011)

PCT

(10) 国際公開番号
WO 2011/114932 A1

- (51) 国際特許分類:
G10L 19/00 (2006.01) H04S 3/00 (2006.01)
- (21) 国際出願番号: PCT/JP2011/055293
- (22) 国際出願日: 2011 年 3 月 8 日(08.03.2011)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2010-061170 2010 年 3 月 17 日(17.03.2010) JP
- (71) 出願人 (米国を除く全ての指定国について): ソニー株式会社(SONY CORPORATION) [JP/JP]; 〒1080075 東京都港区港南 1 丁目 7 番 1 号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 戸栗 康裕 (TOGURI Yasuhiro) [JP/JP]; 〒1080075 東京都港区港南 1 丁目 7 番 1 号 ソニー株式会社内 Tokyo (JP). 鈴木 志朗(SUZUKI Shiro) [JP/JP]; 〒1080075 東京都港区港南 1 丁目 7 番 1 号 ソニー株式会社内 Tokyo (JP). 松本 淳(MATSUMOTO Jun) [JP/JP]; 〒1080075 東京都港区港南 1 丁目 7 番 1 号 ソニー株式会社内 Tokyo (JP). 前田 祐児(MAEDA Yuuji) [JP/JP]; 〒1080075 東京都港区港

南 1 丁目 7 番 1 号 ソニー株式会社内 Tokyo (JP). 松村 祐樹(MATSUMURA Yuuki) [JP/JP]; 〒1080075 東京都港区港南 1 丁目 7 番 1 号 ソニー株式会社内 Tokyo (JP).

- (74) 代理人: 稲本 義雄, 外(INAMOTO Yoshio et al.); 〒1600023 東京都新宿区西新宿 7 丁目 1 1 番 1 8 号 711ビルディング 4 階 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT,

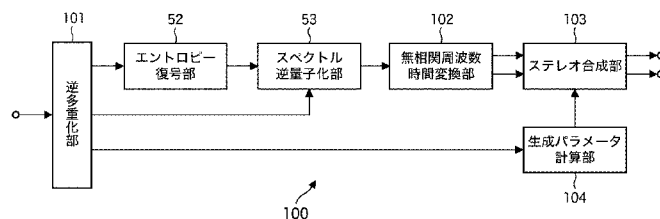
[続葉有]

(54) Title: AUDIO-PROCESSING DEVICE, AUDIO-PROCESSING METHOD AND PROGRAM

(54) 発明の名称: 音声処理装置、音声処理方法、およびプログラム

〔図9〕

図9



- 101 Inverse Multiplexer
52 Entropy Decoder
53 Spectral Inverse Quantizing Unit
102 Non-correlated Frequency and Time Conversion Unit
103 Stereo Synthesizing Unit
104 Generated Parameter Calculator

(57) Abstract: Disclosed is an audio-processing apparatus, an audio-processing method and program that suppress a delay or an increased computing amount when decoding when an audio signal, if a multi-channel audio signal is down-mixed and encoded. An inverse multiplexing unit (101) obtains encoded data multiplexed by a BC parameter. A non-correlated frequency and time conversion unit (102) IMDCT-transforms and IMDST-transforms a frequency spectral coefficient of a monaural signal (X_M) obtained from the encoded data to generate a monaural signal (X_M) that is a time-region signal, and a signal (X_D) that is non-correlated to that monaural signal (X_M). A stereo-synthesizing unit (103) generates a stereo signal by synthesizing the monaural signal (X_M) and signal (X_D) using the BC parameter. This can be applied, for example, to an audio-processing device that decodes down-mixed and encoded stereo signals.

(57) 要約:

[続葉有]

WO 2011/114932 A1



NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI 添付公開書類:

(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, — 国際調査報告 (条約第 21 条(3))
NE, SN, TD, TG).

本発明は、マルチチャンネルのオーディオ信号がダウンミックスされて符号化されている場合に、そのオーディオ信号の復号時の遅延や演算量の増加を抑制することができる音声処理装置、音声処理方法、およびプログラムに関する。逆多重化部 101 は、BC パラメータが多重化された符号化データを取得する。無相関周波数時間変換部 102 は、その符号化データから得られるモノラル信号 X_M の周波数スペクトル係数を IMDCT 変換するとともに IMDST 変換し、時間領域信号であるモノラル信号 X_M と、そのモノラル信号 X_M と略無相関な信号 X_D を生成する。ステレオ合成部 103 は、BC パラメータを用いてモノラル信号 X_M と信号 X_D を合成することにより、ステレオ信号を生成する。本発明は、例えば、ダウンミックスされて符号化されているステレオ信号を復号する音声処理装置に適用することができる。

明 細 書

発明の名称： 音声処理装置、音声処理方法、およびプログラム 技術分野

[0001] 本発明は、音声処理装置、音声処理方法、およびプログラムに関し、特に、マルチチャンネルのオーディオ信号がダウンミックスされて符号化されている場合に、そのオーディオ信号の復号時の遅延や演算量の増加を抑制することができるようにした音声処理装置、音声処理方法、およびプログラムに関する。

背景技術

[0002] マルチチャンネルのオーディオ信号を符号化する符号化装置は、チャンネル間の関係を利用した符号化を行うことで効率の高い符号化を行うことができる。このような符号化としては、例えば、インテンシティ符号化、M/Sステレオ符号化、空間符号化などがある。空間符号化を行う符号化装置は、 n チャンネルのオーディオ信号を m チャンネル ($m < n$) のオーディオ信号にダウンミックスして符号化し、そのダウンミックスの際にチャンネル間の関係を表す空間パラメータを求め、その空間パラメータを符号化データとともに伝送する。空間パラメータと符号化データを受信する復号装置は、符号化データを復号し、空間パラメータを用いて、復号の結果得られる m チャンネルのオーディオ信号から元の n チャンネルのオーディオ信号を復元する。

[0003] このような空間符号化は、バイノーラルキュー符号化(Binaural cue coding)として知られている。空間パラメータ（以下、BCパラメータという）としては、ILD (Inter-channel Level Difference) (チャンネル間レベル差)、IPD (Inter-channel Phase Difference) (チャンネル間位相差)、ICC (Inter-channel Correlation) (チャンネル間相関)などが用いられる。ILDは、チャンネル間の信号の大きさの比率を示すパラメータである。IPDは、チャンネル間の位相差を示すパラメータであり、ICCは、チャンネル間の相関性を示すパラメータである。

- [0004] 図1は、空間符号化を行う符号化装置の構成例を示すブロック図である。
- [0005] なお、以下では、説明を簡単にするため、 $n=2$ 、 $m=1$ とする。即ち、符号化対象のオーディオ信号はステレオのオーディオ信号（以下、ステレオ信号という）であり、符号化の結果得られる符号化データはモノラルのオーディオ信号（以下、モノラル信号という）の符号化データである。
- [0006] 図1の符号化装置10は、チャンネルダウンミックス部11、空間パラメータ検出部12、オーディオ信号符号化部13、および多重化部14により構成される。符号化装置10には、左用のオーディオ信号 X_L と右用のオーディオ信号 X_R からなるステレオ信号が符号化対象として入力され、符号化装置10は、モノラル信号の符号化データを出力する。
- [0007] 具体的には、符号化装置10のチャンネルダウンミックス部11は、符号化対象として入力されたステレオ信号をモノラル信号 X_M にダウンミックスする。そして、チャンネルダウンミックス部11は、モノラル信号を空間パラメータ検出部12とオーディオ信号符号化部13に供給する。
- [0008] 空間パラメータ検出部12は、チャンネルダウンミックス部11から供給されるモノラル信号 X_M と、符号化対象として入力されたステレオ信号とに基づいて、BCパラメータを検出し、多重化部14に供給する。
- [0009] オーディオ信号符号化部13は、チャンネルダウンミックス部11から供給されるモノラル信号を符号化し、その結果得られる符号化データを多重化部14に供給する。
- [0010] 多重化部14は、オーディオ信号符号化部13から供給される符号化データと、空間パラメータ検出部12から供給されるBCパラメータを多重化して出力する。
- [0011] 図2は、図1のオーディオ信号符号化部13の構成例を示すブロック図である。
- [0012] なお、図2のオーディオ信号符号化部13の構成は、オーディオ信号符号化部13が例えばMPEG-2 AAC LC (Moving Picture Experts Group phase 2 Advanced Audio Coding Low Complexity) プロファイル方式で符号化を行う場

合の構成である。但し、説明を簡単にするため、図2では構成を簡略化して記載している。

[0013] 図2のオーディオ信号符号化部13は、MDCT (Modified Discrete Cosine Transform) (修正コサイン変換) 部21、スペクトル量子化部22、エントロピー符号化部23、および多重化部24により構成される。

[0014] MDCT部21は、チャンネルダウンミックス部11から供給されるモノラル信号に対してMDCTを行い、時間領域信号であるモノラル信号を周波数領域の係数であるMDCT係数に変換する。MDCT部21は、変換の結果得られるMDCT係数を周波数スペクトル係数としてスペクトル量子化部22に供給する。

[0015] スペクトル量子化部22は、MDCT部21から供給される周波数スペクトル係数を量子化し、エントロピー符号化部23に供給する。また、スペクトル量子化部22は、この量子化に関する情報である量子化情報を多重化部24に供給する。量子化情報としては、スケールファクタ、量子化ビット情報などがある。

[0016] エントロピー符号化部23は、スペクトル量子化部22から供給される量子化された周波数スペクトル係数に対して、ハフマン符号化、算術符号化などのエントロピー符号化を行い、可逆圧縮する。エントロピー符号化部23は、エントロピー符号化の結果得られるデータを多重化部24に供給する。

[0017] 多重化部24は、エントロピー符号化部23から供給されるデータと、スペクトル量子化部22から供給される量子化情報とを多重化し、その結果得られるデータを符号化データとして多重化部14 (図1) に供給する。

[0018] 図3は、図1のオーディオ信号符号化部13の他の構成例を示すブロック図である。

[0019] なお、図3のオーディオ信号符号化部13の構成は、例えばMPEG-2 AAC SS R (Scalable Sample Rate) プロファイルや、MP3 (MPEG Audio Layer-3) などの方式で符号化を行う場合の構成である。但し、説明を簡単にするため、図3では構成を簡略化して記載している。

[0020] 図3のオーディオ信号符号化部13は、分析フィルタバンク31、MDCT部

32-1乃至32-N（Nは任意の整数）、スペクトル量子化部33、エントロピー符号化部34、および多重化部35により構成される。

[0021] 分析フィルタバンク31は、QMF（Quadrature Mirror Filterbank）バンクや、PQF（Poly-phase Quadrature Filter）バンクなどにより構成される。分析フィルタバンク31は、チャンネルダウンミックス部11から供給されるモノラル信号を周波数によってN個のグループに分割する。分析フィルタバンク31は、分割の結果得られるN個のサブバンド信号をMDCT部32-1乃至32-Nにそれぞれ供給する。

[0022] MDCT部32-1乃至32-Nは、それぞれ、分析フィルタバンク31から供給されるサブバンド信号に対してMDCTを行い、時間領域信号であるサブバンド信号を周波数領域の係数であるMDCT係数に変換する。そして、MDCT部32-1乃至32-Nは、それぞれ、各サブバンド信号のMDCT係数を周波数スペクトル係数として、スペクトル量子化部33に供給する。

[0023] スペクトル量子化部33は、MDCT部32-1乃至32-Nから供給されるN個の周波数スペクトル係数を、それぞれ、量子化し、エントロピー符号化部34に供給する。また、スペクトル量子化部33は、この量子化の量子化情報を多重化部35に供給する。

[0024] エントロピー符号化部34は、スペクトル量子化部33から供給されるN個の量子化された周波数スペクトル係数のそれぞれに対して、ハフマン符号化や算術符号化などのエントロピー符号化を行い、可逆圧縮する。エントロピー符号化部34は、エントロピー符号化の結果得られるN個のデータを多重化部35に供給する。

[0025] 多重化部35は、エントロピー符号化部34から供給されるN個のデータと、スペクトル量子化部33から供給される量子化情報とを多重化し、その結果得られるデータを符号化データとして多重化部14（図1）に供給する。

[0026] 図4は、図1の符号化装置10により空間符号化された符号化データを復号する復号装置の構成例を示すブロック図である。

- [0027] 図4の復号装置40は、逆多重化部41、オーディオ信号復号部42、生成パラメータ計算部43、およびステレオ信号生成部44により構成される。復号装置40は、図1の符号化装置から供給される符号化データを復号し、ステレオ信号を生成する。
- [0028] 具体的には、復号装置40の逆多重化部41は、図1の符号化装置10から供給される多重化された符号化データに対して逆多重化を行い、符号化データとBCパラメータを得る。逆多重化部41は、符号化データをオーディオ信号復号部42に供給し、BCパラメータを生成パラメータ計算部43に供給する。
- [0029] オーディオ信号復号部42は、逆多重化部41から供給される符号化データを復号し、その結果得られる時間領域信号であるモノラル信号 X_M をステレオ信号生成部44に供給する。
- [0030] 生成パラメータ計算部43は、逆多重化部41から供給されるBCパラメータを用いて、それとともに多重化された符号化データの復号結果であるモノラル信号からステレオ信号を生成するためのパラメータである生成パラメータを計算する。生成パラメータ計算部43は、その生成パラメータをステレオ信号生成部44に供給する。
- [0031] ステレオ信号生成部44は、生成パラメータ計算部43から供給される生成パラメータを用いて、オーディオ信号復号部42から供給されるモノラル信号 X_M から、左用のオーディオ信号 X_L と右用のオーディオ信号 X_R を生成する。ステレオ信号生成部44は、その左用のオーディオ信号 X_L と右用のオーディオ信号 X_R をステレオ信号として出力する。
- [0032] 図5は、図4のオーディオ信号復号部42の構成例を示すブロック図である。
- [0033] なお、図5のオーディオ信号復号部42の構成は、例えばMPEG-2 AAC LCプロファイル方式で符号化された符号化データが復号装置40に入力される場合の構成である。即ち、図5のオーディオ信号復号部42は、図2のオーディオ信号符号化部13で符号化された符号化データを復号するものである。

- [0034] 図5のオーディオ信号復号部42は、逆多重化部51、エントロピー復号部52、スペクトル逆量子化部53、およびIMDCT部54により構成される。
- [0035] 逆多重化部51は、図4の逆多重化部41から供給される符号化データに対して逆多重化を行い、量子化され、エントロピー符号化された周波数スペクトル係数と量子化情報を得る。逆多重化部51は、量子化され、エントロピー符号化された周波数スペクトル係数をエントロピー復号部52に供給し、量子化情報をスペクトル逆量子化部53に供給する。
- [0036] エントロピー復号部52は、逆多重化部51から供給される周波数スペクトル係数に対して、ハフマン復号や算術復号などのエントロピー復号を行い、量子化された周波数スペクトル係数を復元する。エントロピー復号部52は、その周波数スペクトル係数をスペクトル逆量子化部53に供給する。
- [0037] スペクトル逆量子化部53は、逆多重化部51から供給される量子化情報に基づいて、エントロピー復号部52から供給される量子化された周波数スペクトル係数を逆量子化し、周波数スペクトル係数を復元する。そして、スペクトル逆量子化部53は、その周波数スペクトル係数をIMDCT（Inverse MDCT）（逆修正コサイン変換）部54に供給する。
- [0038] IMDCT部54は、スペクトル逆量子化部53から供給される周波数スペクトル係数に対してIMDCTを行い、周波数スペクトル係数を時間領域信号であるモノラル信号 X_M に変換する。IMDCT部54は、そのモノラル信号 X_M をステレオ信号生成部44（図4）に供給する。
- [0039] 図6は、図4のオーディオ信号復号部42の他の構成例を示すブロック図である。
- [0040] なお、図6のオーディオ信号復号部42の構成は、例えばMPEG-2 AAC SSRプロファイルや、MP3などの方式で符号化された符号化データが復号装置40に入力される場合の構成である。即ち、図6のオーディオ信号復号部42は、図3のオーディオ信号符号化部13で符号化された符号化データを復号するものである。
- [0041] 図6のオーディオ信号復号部42は、逆多重化部61、エントロピー復号

部 6 2、スペクトル逆量子化部 6 3、IMDCT 部 6 4 - 1 乃至 6 4 - N、および合成フィルタバンク 6 5 により構成される。

[0042] 逆多重化部 6 1 は、図 4 の逆多重化部 4 1 から供給される符号化データに対して逆多重化を行い、量子化され、エントロピー符号化された N 個のサブバンド信号の周波数スペクトル係数と量子化情報を得る。逆多重化部 6 1 は、量子化され、エントロピー符号化された N 個のサブバンド信号の周波数スペクトル係数をエントロピー復号部 6 2 に供給し、量子化情報をスペクトル逆量子化部 6 3 に供給する。

[0043] エントロピー復号部 6 2 は、逆多重化部 6 1 から供給される N 個のサブバンド信号の周波数スペクトル係数のそれぞれに対してハフマン復号や算術復号などのエントロピー復号を行い、スペクトル逆量子化部 6 3 に供給する。

[0044] スペクトル逆量子化部 6 3 は、逆多重化部 6 1 から供給される量子化情報に基づいて、エントロピー復号部 6 2 から供給されるエントロピー復号の結果得られた N 個のサブバンド信号の周波数スペクトル係数をそれぞれ逆量子化する。これにより、N 個のサブバンド信号の周波数スペクトル係数が復元される。スペクトル逆量子化部 6 3 は、復元された N 個のサブバンド信号の周波数スペクトル係数を IMDCT 部 6 4 - 1 乃至 6 4 - N に 1 つずつ供給する。

[0045] IMDCT 部 6 4 - 1 乃至 6 4 - N は、それぞれ、スペクトル逆量子化部 6 3 から供給される周波数スペクトル係数に対して IMDCT を行い、周波数スペクトル係数を時間領域信号であるサブバンド信号に変換する。IMDCT 部 6 4 - 1 乃至 6 4 - N は、変換の結果得られるサブバンド信号を、それぞれ合成フィルタバンク 6 5 に供給する。

[0046] 合成フィルタバンク 6 5 は、逆PQFや逆QMFなどにより構成される。合成フィルタバンク 6 5 は、IMDCT 部 6 4 - 1 乃至 6 4 - N から供給される N 個のサブバンド信号を合成し、その結果得られる信号をモノラル信号 X_M としてステレオ信号生成部 4 4（図 4）に供給する。

[0047] 図 7 は、図 4 のステレオ信号生成部 4 4 の構成例を示すブロック図である。

- [0048] 図7のステレオ信号生成部44は、残響信号生成部71とステレオ合成部72により構成される。
- [0049] 残響信号生成部71は、図4のオーディオ信号復号部42から供給されるモノラル信号 X_M を用いて、このモノラル信号 X_M とは無相関な信号 X_D を生成する。残響信号生成部71としては、一般的に、コムフィルタやオールパスフィルタなどが用いられる。この場合、残響信号生成部71は、モノラル信号 X_M の残響(リバーブ)信号を信号 X_D として生成する。
- [0050] なお、残響信号生成部71としては、フィードバック遅延ネットワーク(Feedback Delay Network(FDN))が用いられることもある(例えば、特許文献1参照)。
- [0051] 残響信号生成部71は、生成された信号 X_D をステレオ合成部72に供給する。
- [0052] ステレオ合成部72は、図4の生成パラメータ計算部43から供給される生成パラメータを用いて、図4のオーディオ信号復号部42から供給されるモノラル信号 X_M と、残響信号生成部71から供給される信号 X_D とを合成する。そして、ステレオ合成部72は、合成の結果得られる左用のオーディオ信号 X_L と右用のオーディオ信号 X_R をステレオ信号として出力する。
- [0053] 図8は、図4のステレオ信号生成部44の他の構成例を示すブロック図である。
- [0054] 図8のステレオ信号生成部44は、分析フィルタバンク81、サブバンドステレオ信号生成部82-1乃至82-P(Pは任意の数)、および合成フィルタバンク83により構成される。
- [0055] なお、図4のステレオ信号生成部44の構成が図8に示す構成である場合、図1の符号化装置10の空間パラメータ検出部12では、サブバンド信号ごとにBCパラメータが検出される。
- [0056] 具体的には、例えば、空間パラメータ検出部12は、2つの分析フィルタバンクを有する。そして、空間パラメータ検出部12は、一方の分析フィルタバンクでステレオ信号を周波数によって分割し、他方の分析フィルタバン

クでチャンネルダウンミックス部 11 からのモノラル信号を周波数によって分割する。空間パラメータ検出部 12 は、分割の結果得られるステレオ信号のサブバンド信号とモノラル信号のサブバンド信号に基づいて、サブバンド信号ごとに BC パラメータを検出する。そして、図 4 の生成パラメータ計算部 43 には、逆多重化部 41 から各サブバンド信号の BC パラメータが供給され、生成パラメータ計算部 43 は、サブバンド信号ごとに生成パラメータを生成する。

[0057] 分析フィルタバンク 81 は、QMF (Quadrature Mirror Filter) バンクなどにより構成される。分析フィルタバンク 81 は、図 4 のオーディオ信号復号部 42 から供給されるモノラル信号 X_M を周波数によって P 個のグループに分割する。分析フィルタバンク 81 は、分割の結果得られる P 個のサブバンド信号を、サブバンドステレオ信号生成部 82-1 乃至 82- P にそれぞれ供給する。

[0058] サブバンドステレオ信号生成部 82-1 乃至 82- P は、それぞれ、残響信号生成部とステレオ合成部により構成される。各サブバンドステレオ信号生成部 82-1 乃至 82- P の構成は同一であるので、ここでは、サブバンドステレオ信号生成部 82-B についてのみ説明する。

[0059] サブバンドステレオ信号生成部 82-B は、残響信号生成部 91 とステレオ合成部 92 により構成される。残響信号生成部 91 は、分析フィルタバンク 81 から供給されるモノラル信号のサブバンド信号 X_m^B を用いて、このサブバンド信号 X_m^B とは無関係な信号 X_D^B を生成し、信号 X_D^B をステレオ合成部 92 に供給する。

[0060] ステレオ合成部 92 は、図 4 の生成パラメータ計算部 43 から供給されるサブバンド信号 X_m^B の生成パラメータを用いて、分析フィルタバンク 81 から供給されるサブバンド信号 X_m^B と、残響信号生成部 91 から供給される信号 X_D^B とを合成する。そして、ステレオ合成部 92 は、合成の結果得られる左用のオーディオ信号 X_L^B と右用のオーディオ信号 X_R^B を、ステレオ信号のサブバンド信号として合成フィルタバンク 83 に供給する。

- [0061] 合成フィルタバンク 8 3 は、サブバンドステレオ信号生成部 8 2 - 1 乃至 8 2 - P から供給される各サブバンド信号のステレオ信号を左用および右用ごとに合成する。合成フィルタバンク 8 3 は、その結果得られる左用のオーディオ信号 X_L と右用のオーディオ信号 X_R をステレオ信号として出力する。
- [0062] なお、図 8 のステレオ信号生成部 4 4 の構成は、例えば、特許文献 2 に記載されている。
- [0063] また、インテンシティ符号化を行う符号化装置は、入力されたステレオ信号の所定の周波数帯域以上の周波数の各チャンネルの周波数スペクトル係数をミックスし、モノラル信号の周波数スペクトル係数を生成する。そして、符号化装置は、このモノラル信号の周波数スペクトル係数、および、チャンネル間の周波数スペクトル係数のレベル比を符号化結果として出力する。
- [0064] 具体的には、インテンシティ符号化を行う符号化装置は、ステレオ信号に対して MDCT 変換を行い、その結果得られる各チャンネルの周波数スペクトル係数のうち、所定の周波数帯域以上の周波数の各チャンネルの周波数スペクトル係数をミックスして共通化する。そして、インテンシティ符号化を行う符号化装置は、共通化された周波数スペクトル係数を量子化してエントロピー符号化を行い、その結果得られるデータを量子化情報と多重化して符号化データとする。また、インテンシティ符号化を行う符号化装置は、チャンネル間の周波数スペクトル係数のレベル比を求め、そのレベル比を符号化データと多重化して出力する。
- [0065] また、インテンシティ復号を行う復号装置は、チャンネル間の周波数スペクトル係数のレベル比が多重化された符号化データに対して逆多重化を行い、その結果得られる符号化データをエントロピー復号し、量子化情報に基づいて逆量子化する。また、インテンシティ復号を行う復号装置は、逆量子化の結果得られた周波数スペクトル係数と、符号化データに多重化されたチャンネル間の周波数スペクトル係数のレベル比とに基づいて、各チャンネルの周波数スペクトル係数を復元する。そして、インテンシティ復号を行う復号装置は、復元された各チャンネルの周波数スペクトル係数に対して IMDCT を行

い、所定の周波数帯域以上の周波数のステレオ信号を得る。

- [0066] このようなインテンシティ符号化は、符号化効率を向上させるためによく用いられるが、ステレオ信号の高域の周波数スペクトル係数をモノラル化してチャンネル間のレベル差のみで表現しているため、本来のステレオ感がやや失われる。

先行技術文献

特許文献

- [0067] 特許文献1：特開2006-325162号公報

特許文献2：特表2006-524832号公報

発明の概要

発明が解決しようとする課題

- [0068] 上述したように、従来の空間符号化された符号化データを復号する復号装置40は、ステレオ信号の生成の際に用いられるモノラル信号 X_M と無関係な信号 X_D や信号 X_{D^1} 乃至 X_{D^P} を、時間領域信号であるモノラル信号 X_M を用いて生成する。
- [0069] 従って、信号 X_D を生成する残響信号生成部71や、信号 X_{D^1} 乃至 X_{D^P} を生成する分析フィルタバンク81とサブバンドステレオ信号生成部82-1乃至82-Pの残響信号生成部91によって遅延が発生し、復号装置40のアルゴリズム遅延が増大する。このことは、例えば、復号装置40に即時の応答特性が要求される場合や復号装置40がリアルタイム通信に用いられる場合などの低遅延特性が重要になる場合に問題となる。
- [0070] また、残響信号生成部71や、分析フィルタバンク81とサブバンドステレオ信号生成部82-1乃至82-Pの残響信号生成部91におけるフィルタ演算などにより、演算量が増大し、必要なバッファ容量も増大する。
- [0071] 本発明は、このような状況に鑑みてなされたものであり、マルチチャンネルのオーディオ信号がダウンミックスされて符号化されている場合に、そのオーディオ信号の復号時の遅延や演算量の増加を抑制することができるよう

にするものである。

課題を解決するための手段

[0072] 本発明の一側面の音声処理装置は、複数チャンネルの音声の時間領域信号である音声信号から生成された前記複数チャンネルより少ないチャンネルの音声信号の周波数領域の係数と、前記複数チャンネルのチャンネル間の関係を表すパラメータとを取得する取得手段と、前記取得手段により取得された前記周波数領域の係数を、第1の時間領域信号に変換する第1の変換手段と、前記取得手段により取得された前記周波数領域の係数を、第2の時間領域信号に変換する第2の変換手段と、前記パラメータを用いて前記第1の時間領域信号と前記第2の時間領域信号を合成することにより、前記複数チャンネルの音声信号を生成する合成手段とを備え、前記第1の変換手段による変換における基底と前記第2の変換手段による変換における基底は直交する音声処理装置である。

[0073] 本発明の一側面の音声処理方法およびプログラムは、本発明の一側面の音声処理装置に対応する。

[0074] 本発明の一側面においては、複数チャンネルの音声の時間領域信号である音声信号から生成された前記複数チャンネルより少ないチャンネルの音声信号の周波数領域の係数と、前記複数チャンネルのチャンネル間の関係を表すパラメータとが取得され、取得された前記周波数領域の係数が、第1の時間領域信号に変換され、取得された前記周波数領域の係数が、第2の時間領域信号に変換され、前記パラメータを用いて前記第1の時間領域信号と前記第2の時間領域信号が合成されることにより、前記複数チャンネルの音声信号が生成される。なお、第1の時間領域信号への変換における基底と第2の時間領域信号への変換における基底は直交する。

[0075] 本発明の一側面の音声処理装置は、独立した装置であっても良いし、1つの装置を構成している内部ブロックであっても良い。

発明の効果

[0076] 本発明の一側面によれば、マルチチャンネルのオーディオ信号がダウンミ

ックスされて符号化されている場合に、そのオーディオ信号の復号時の遅延や演算量の増加を抑制することができる。

図面の簡単な説明

- [0077] [図1]空間符号化を行う符号化装置の構成例を示すブロック図である。
- [図2]図1のオーディオ信号符号化部の構成例を示すブロック図である。
- [図3]図1のオーディオ信号符号化部の他の構成例を示すブロック図である。
- [図4]空間符号化された符号化データを復号する復号装置の構成例を示すブロック図である。
- [図5]図4のオーディオ信号復号部の構成例を示すブロック図である。
- [図6]図4のオーディオ信号復号部の他の構成例を示すブロック図である。
- [図7]図4のステレオ信号生成部の構成例を示すブロック図である。
- [図8]図4のステレオ信号生成部の他の構成例を示すブロック図である。
- [図9]本発明を適用した音声処理装置の第1実施の形態の構成例を示すブロック図である。
- [図10]図9の無相関周波数時間変換部の詳細構成例を示すブロック図である。
- [図11]図9の無相関周波数時間変換部の他の詳細構成例を示すブロック図である。
- [図12]図9のステレオ合成部の詳細構成例を示すブロック図である。
- [図13]各信号のベクトルを表す図である。
- [図14]図9の音声処理装置による復号処理を説明するフローチャートである。
- [図15]本発明を適用した音声処理装置の第2実施の形態の構成例を示すブロック図である。
- [図16]図15の音声処理装置による復号処理を説明するフローチャートである。
- [図17]本発明を適用した音声処理装置の第3実施の形態の構成例を示すブロック図である。

[図18] 図 17 の音声処理装置による復号処理を説明するフローチャートである。

[図19] 本発明を適用した音声処理装置の第 4 実施の形態の構成例を示すブロック図である。

[図20] 図 19 の音声処理装置による復号処理を説明するフローチャートである。

[図21] コンピュータの一実施の形態の構成例を示す図である。

発明を実施するための形態

[0078] <第 1 実施の形態>

[音声処理装置の第 1 実施の形態の構成例]

図 9 は、本発明を適用した音声処理装置の第 1 実施の形態の構成例を示すブロック図である。

[0079] 図 9 に示す構成のうち、図 4 および図 5 の構成と同じ構成には同じ符号を付してある。重複する説明については適宜省略する。

[0080] 図 9 の音声処理装置 100 の構成は、主に、逆多重化部 41 および逆多重化部 51 の代わりに逆多重化部 101 が設けられている点、IMDCT 部 54 および残響信号生成部 71 の代わりに無相関周波数時間変換部 102 が設けられている点、および、ステレオ合成部 72、生成パラメータ計算部 43 の代わりに、ステレオ合成部 103、生成パラメータ計算部 104 が設けられている点が、図 5 のオーディオ信号復号部 42 と図 7 のステレオ信号生成部 44 を備える図 4 の復号装置 40 の構成と異なる。

[0081] 音声処理装置 100 は、例えば、図 2 のオーディオ信号符号化部 13 を備える図 1 の符号化装置 10 により空間符号化された符号化データを復号する。このとき、音声処理装置 100 は、ステレオ信号の生成の際に用いられるモノラル信号 X_M と無関係な信号 X_D' を、モノラル信号 X_M の周波数スペクトル係数を用いて生成する。

[0082] 具体的には、音声処理装置 100 の逆多重化部 101（取得手段）は、図 4 の逆多重化部 41 と図 5 の逆多重化部 51 に対応する。即ち、逆多重化部

101は、図1の符号化装置10から供給される多重化された符号化データに対して逆多重化を行い、符号化データとBCパラメータを取得する。なお、符号化データに多重化されるBCパラメータは、全てのフレームについてのBCパラメータであってもよいし、所定のフレームについてのBCパラメータであってもよいが、ここでは、所定のフレームについてのBCパラメータであるものとする。

[0083] また、逆多重化部101は、符号化データに対して逆多重化を行い、量子化され、エントロピー符号化された周波数スペクトル係数と量子化情報を得る。そして、逆多重化部101は、量子化され、エントロピー符号化された周波数スペクトル係数をエントロピー復号部52に供給し、量子化情報をスペクトル逆量子化部53に供給する。また、逆多重化部101は、BCパラメータを生成パラメータ計算部104に供給する。

[0084] 無相関周波数時間変換部102は、スペクトル逆量子化部53による逆量子化の結果得られるモノラル信号 X_M の周波数スペクトル係数から、互いに無相関な2つの時間領域信号であるモノラル信号 X_M と信号 X_D' を生成する。そして、無相関周波数時間変換部102は、モノラル信号 X_M と信号 X_D' をステレオ合成部103に供給する。この無相関周波数時間変換部102の詳細は、後述する図10や図11を参照して説明する。

[0085] ステレオ合成部103（合成手段）は、生成パラメータ計算部104から供給される生成パラメータを用いて、無相関周波数時間変換部102から供給されるモノラル信号 X_M と信号 X_D' とを合成する。そして、ステレオ合成部103は、合成の結果得られる左用のオーディオ信号 X_L と右用のオーディオ信号 X_R をステレオ信号として出力する。このステレオ合成部103の詳細は、後述する図12を参照して説明する。

[0086] 生成パラメータ計算部104は、逆多重化部101から供給される所定のフレームについてのBCパラメータを補間し、各フレームのBCパラメータを計算する。生成パラメータ計算部104は、現在の処理対象のフレームのBCパラメータを用いて生成パラメータを生成し、ステレオ合成部103に

供給する。

[0087] [無相関周波数時間変換部の詳細構成例]

図 10 は、図 9 の無相関周波数時間変換部 102 の詳細構成例を示すブロック図である。

[0088] 図 10 の無相関周波数時間変換部 102 は、IMDCT 部 54 と IMDST 部 111 により構成される。

[0089] 図 10 の IMDCT 部 54 (第 1 の変換手段) は、図 5 の IMDCT 部 54 と同一のものであり、スペクトル逆量子化部 53 から供給されるモノラル信号 X_M の周波数スペクトル係数に対して IMDCT を行う。そして、IMDCT 部 54 は、その結果得られる時間領域信号であるモノラル信号 X_M (第 1 の時間領域信号) をステレオ合成部 103 (図 9) に供給する。

[0090] IMDST (Inverse Modified Discrete Sine Transform) 部 111 (第 2 の変換手段) は、スペクトル逆量子化部 53 から供給されるモノラル信号 X_M の周波数スペクトル係数に対して IMDST を行う。そして、IMDST 部 111 は、その結果得られる時間領域信号である信号 X_D' (第 2 の時間領域信号) をステレオ合成部 103 (図 9) に供給する。

[0091] 以上のように、IMDCT 部 54 による変換はコサインの逆変換であり、IMDST 部 111 による変換はサインの逆変換であり、IMDCT 部 54 による変換における基底と IMDST 部 111 による変換における基底は直交している。従って、モノラル信号 X_M と信号 X_D' は、互いに略無相関な信号とみなすことができる。

[0092] なお、MDCT、IMDCT、および IMDST は、それぞれ、以下の式 (1) 乃至 (3) で定義される。

[0093] [数 1]

$$X_c(k) = \sum_{n=0}^{2N-1} w(n) \cdot x(n) \cdot \cos \left[\frac{\pi}{4N} (2n+1+N)(2k+1) \right] \quad k=0,1,\dots,N-1 \quad \cdots (1)$$

[数2]

$$y(n) = \frac{2 \cdot w'(n)}{N} \cdot \sum_{k=0}^{N-1} X_c(k) \cdot \cos \left[\frac{\pi}{4N} (2n+1+N)(2k+1) \right] \quad n=0,1,\dots,2N-1 \quad \cdots(2)$$

[数3]

$$y(n) = \frac{2 \cdot w'(n)}{N} \cdot \sum_{k=0}^{N-1} X_s(k) \cdot \sin \left[\frac{\pi}{4N} (2n+1+N)(2k+1) \right] \quad n=0,1,\dots,2N-1 \quad \cdots(3)$$

[0094] 式(1)乃至(3)において、 $x(n)$ は時間領域信号であり、 $w(n)$ は変換窓であり、 $w'(n)$ は逆変換窓であり、 $y(n)$ は逆変換後の信号である。また、 $X_c(k)$ はMDCT係数であり、 $X_s(k)$ はMDST係数である。

[0095] [無相関周波数時間変換部の詳細構成例]

図11は、図9の無相関周波数時間変換部102の他の詳細構成例を示すブロック図である。

[0096] 図11に示す構成のうち、図10の構成と同じ構成には同じ符号を付してある。重複する説明については適宜省略する。

[0097] 図11の無相関周波数時間変換部102の構成は、主に、IMDST部111の代わりにスペクトル反転部121、IMDCT部122、および符号反転部123が設けられている点が図10の構成と異なる。

[0098] 図11の無相関周波数時間変換部102のスペクトル反転部121は、スペクトル逆量子化部53から供給される周波数スペクトル係数を、周波数が逆順になるように反転し、IMDCT部122に供給する。

[0099] IMDCT部122は、スペクトル反転部121から供給される周波数スペクトル係数に対してIMDCTを行い、時間領域信号を得る。IMDCT部122は、その時間領域信号を符号反転部123に供給する。

[0100] 符号反転部123は、IMDCT部122から供給される時間領域信号の奇数サンプルの符号を反転し、信号 x_D' を得る。

[0101] ここで、IMDSTを定義する上述した式(3)において、 $X_s(k)$ を $X_s(N-k-1)$ に置き換えると、 N が一般的な4の倍数とすれば、式(3)は、以下の式(4)

に変形することができる。

[0102] [数4]

$$\begin{aligned}
 y(n) &= \frac{2 \cdot w'(n)}{N} \cdot \sum_{k=0}^{N-1} X_s(N-k-1) \cdot \sin \left[\frac{\pi}{4N} (2n+1+N)(2(N-k-1)+1) \right] \\
 &= \frac{2 \cdot w'(n)}{N} \cdot (-1)^n \cdot \sum_{k=0}^{N-1} X_s(N-k-1) \cdot \cos \left[\frac{\pi}{4N} (2n+1+N)(2k+1) \right] = (-1)^n \cdot \text{IMDCT}[X_s(N-k-1)] \\
 &\quad \dots(4)
 \end{aligned}$$

[0103] 従って、スペクトル逆量子化部 53 からの周波数スペクトル係数に対して IMDST を行った結果得られる信号と、その周波数スペクトル係数を周波数が逆順になるように反転して IMDST を行い、奇数サンプルの符号を反転した結果得られる信号は、同一の信号 X_D' となる。即ち、図 10 の IMDST 部 111 と、図 11 のスペクトル反転部 121、IMDCT 部 122、および符号反転部 123 とは、等価である。

[0104] 符号反転部 123 は、得られた信号 X_D' を図 9 のステレオ合成部 103 に供給する。

[0105] 以上のように、図 11 の無相関周波数時間変換部 102 は、時間領域信号を周波数スペクトル係数に変換するために IMDCT 部だけを設ければよいので、図 9 の IMDCT 部と IMDST 部を設ける必要がある場合に比べて、製造コストを削減することができる。

[0106] [ステレオ合成部の詳細構成例]

図 12 は、図 9 のステレオ合成部 103 の詳細構成例を示すブロック図である。

[0107] 図 12 のステレオ合成部 103 は、乗算器 141 乃至 144 並びに加算器 145 および加算器 146 により構成される。

[0108] 乗算器 141 は、無相関周波数時間変換部 102 から供給されるモノラル信号 X_M に対して、生成パラメータ計算部 104 から供給される生成パラメータの 1 つである係数 h_{11} を乗算する。乗算器 141 は、その結果得られる乗算値 $h_{11} \times X_M$ を加算器 145 に供給する。

- [0109] 乗算器 142 は、無相関周波数時間変換部 102 から供給されるモノラル信号 X_M に対して、生成パラメータ計算部 104 から供給される生成パラメータの 1 つである係数 h_{21} を乗算する。乗算器 141 は、その結果得られる乗算値 $h_{21} \times X_M$ を加算器 146 に供給する。
- [0110] 乗算器 143 は、無相関周波数時間変換部 102 から供給される信号 X_D' に対して、生成パラメータ計算部 104 から供給される生成パラメータの 1 つである係数 h_{12} を乗算する。乗算器 141 は、その結果得られる乗算値 $h_{12} \times X_D'$ を加算器 145 に供給する。
- [0111] 乗算器 144 は、無相関周波数時間変換部 102 から供給される信号 X_D' に対して、生成パラメータ計算部 104 から供給される生成パラメータの 1 つである係数 h_{22} を乗算する。乗算器 141 は、その結果得られる乗算値 $h_{22} \times X_D'$ を加算器 146 に供給する。
- [0112] 加算器 145 は、乗算器 141 から供給される乗算値 $h_{11} \times X_M$ と、乗算器 143 から供給される乗算値 $h_{12} \times X_D'$ を加算し、その結果得られる加算値を左用のオーディオ信号 X_L として出力する。
- [0113] 加算器 146 は、乗算器 142 から供給される乗算値 $h_{21} \times X_M$ と、乗算器 144 から供給される乗算値 $h_{22} \times X_D'$ を加算し、その結果得られる加算値を右用のオーディオ信号 X_R として出力する。
- [0114] 以上のように、ステレオ合成部 103 では、図 13 に示すように、モノラル信号 X_M 、信号 X_D' 、左用のオーディオ信号 X_L 、および右用のオーディオ信号 X_R をベクトルとして、以下の式 (5) に示すように、生成パラメータを用いた重み付け加算が行われる。

[0115] [数5]

$$\begin{aligned} X_L &= h_{11} \cdot X_M + h_{12} \cdot X_D' \\ X_R &= h_{21} \cdot X_M + h_{22} \cdot X_D' \end{aligned} \quad \cdots (5)$$

- [0116] なお、係数 h_{11} 、 h_{12} 、 h_{21} 、および h_{22} は、以下の式 (6) で表される。

[0117] [数6]

$$\begin{aligned} h_{11} &= g_L \cdot \cos(\theta_L) \\ h_{12} &= g_L \cdot \sin(\theta_L) \\ h_{21} &= g_R \cdot \cos(\theta_R) \\ h_{22} &= g_R \cdot \sin(\theta_R) \end{aligned} \quad \cdots (6)$$

但し、

[数7]

$$g_L = \frac{\|X_L\|}{\|X_M\|}, \quad g_R = \frac{\|X_R\|}{\|X_M\|} \quad \cdots (7)$$

[0118] 式(6)において、角度 θ_L は、左用のオーディオ信号 X_L のベクトルとモノラル信号 X_M のベクトルがなす角度であり、角度 θ_R は、右用のオーディオ信号 X_R のベクトルとモノラル信号 X_M のベクトルがなす角度である。

[0119] ここで、係数 h_{11} 、 h_{12} 、 h_{21} 、および h_{22} は、生成パラメータ計算部104により生成パラメータとして計算される。具体的には、生成パラメータ計算部104は、BCパラメータから g_L 、 g_R 、 θ_L 、および θ_R を計算し、その g_L 、 g_R 、 θ_L 、および θ_R から係数 h_{11} 、 h_{12} 、 h_{21} 、および h_{22} を計算して生成パラメータとする。なお、BCパラメータから g_L 、 g_R 、 θ_L 、および θ_R を計算する方法の詳細は、例えば、特開2006-325162号公報などに記載されている。

[0120] なお、BCパラメータとしては、 g_L 、 g_R 、 θ_L 、および θ_R を用いることもできるし、 g_L 、 g_R 、 θ_L 、および θ_R を圧縮符号化したものを用いることもできる。また、BCパラメータとしては、係数 h_{11} 、 h_{12} 、 h_{21} 、および h_{22} を直接、または圧縮符号化して用いることもできる。

[0121] [音声処理装置の処理の説明]

図14は、図9の音声処理装置100による復号処理を説明するフローチャートである。この復号処理は、図1の符号化装置10から供給される多重化された符号化データが音声処理装置100に入力されたとき、開始される

- 。
- [0122] 図14のステップS11において、逆多重化部101は、図1の符号化装置10から供給される多重化された符号化データに対して逆多重化を行い、符号化データとBCパラメータを取得する。また、逆多重化部101は、その符号化データに対してさらに逆多重化を行い、量子化され、エントロピー符号化された周波数スペクトル係数と量子化情報を取得する。そして、逆多重化部101は、量子化され、エントロピー符号化された周波数スペクトル係数をエントロピー復号部52に供給し、量子化情報をスペクトル逆量子化部53に供給する。また、逆多重化部101は、BCパラメータを生成パラメータ計算部104に供給する。
- [0123] ステップS12において、エントロピー復号部52は、逆多重化部101から供給される周波数スペクトル係数に対して、ハフマン復号や算術復号などのエントロピー復号を行い、量子化された周波数スペクトル係数を復元する。エントロピー復号部52は、その周波数スペクトル係数をスペクトル逆量子化部53に供給する。
- [0124] ステップS13において、スペクトル逆量子化部53は、逆多重化部101から供給される量子化情報に基づいて、エントロピー復号部52から供給される量子化された周波数スペクトル係数に対して逆量子化を行い、周波数スペクトル係数を復元する。そして、スペクトル逆量子化部53は、その周波数スペクトル係数を無相関周波数時間変換部102に供給する。
- [0125] ステップS14において、無相関周波数時間変換部102は、スペクトル逆量子化部53による逆量子化の結果得られるモノラル信号 X_M の周波数スペクトル係数から、互いに無相関な2つの時間領域信号であるモノラル信号 X_M と信号 X_D' を生成する。そして、無相関周波数時間変換部102は、モノラル信号 X_M と信号 X_D' をステレオ合成部103に供給する。
- [0126] ステップS15において、ステレオ合成部103は、生成パラメータ計算部104から供給される生成パラメータを用いて、無相関周波数時間変換部102から供給されるモノラル信号 X_M と信号 X_D' とを合成する。

- [0127] ステップS 1 6において、生成パラメータ計算部 1 0 4は、逆多重化部 1 0 1から供給される所定のフレームについてのBCパラメータを補間し、各フレームについてのBCパラメータを計算する。
- [0128] ステップS 1 7において、生成パラメータ計算部 1 0 4は、現在の処理対象のフレームのBCパラメータを用いて係数 h_{11} 、 h_{12} 、 h_{21} 、および h_{22} を生成パラメータとして生成し、ステレオ合成部 1 0 3に供給する。
- [0129] ステップS 1 8において、ステレオ合成部 1 0 3は、生成パラメータ計算部 1 0 4から供給される生成パラメータを用いて、無相関周波数時間変換部 1 0 2から供給されるモノラル信号 X_M と信号 X_D' を合成し、ステレオ信号を生成する。そして、ステレオ合成部 1 0 3はステレオ信号を出力し、処理は終了する。
- [0130] 以上のように、音声処理装置 1 0 0は、モノラル信号 X_M の周波数スペクトル係数に対して基底が直交する2つの変換を行うことによりモノラル信号 X_M と信号 X_D' を生成する。即ち、音声処理装置 1 0 0では、モノラル信号 X_M の周波数スペクトル係数を用いて信号 X_D' を生成することができる。従って、音声処理装置 1 0 0では、従来の図5のオーディオ信号復号部42と図7のステレオ信号生成部44を備える図4の復号装置40に比べて、図7の残響信号生成部71による遅延、演算量やバッファなどのリソースの増加を抑制することができる。
- [0131] また、従来の復号装置40のIMDCT部54を無相関周波数時間変換部102の一部に再利用することができるので、新たな機能の追加が最小限で済み、回路規模や必要なリソースの増加を抑制することができる。
- [0132] <第2実施の形態>
- [音声処理装置の第2実施の形態の構成例]
- 図15は、本発明を適用した音声処理装置の第2実施の形態の構成例を示すブロック図である。
- [0133] 図15に示す構成のうち、図9の構成と同じ構成には同じ符号を付してある。重複する説明については適宜省略する。

[0134] 図15の音声処理装置200の構成は、主に、新たに帯域分割部201、IMDCT部202、加算器203、および加算器204が設けられている点が図9の構成と異なる。

[0135] 音声処理装置200は、例えば、図2のオーディオ信号符号化部13を備える図1の符号化装置10と同様の空間符号化が行われ、高域についてのBCパラメータが多重化された符号化データを復号し、高域のモノラル信号 X_M のみをステレオ化する。

[0136] 具体的には、音声処理装置200の帯域分割部201（分割手段）は、スペクトル逆量子化部53により得られる周波数スペクトル係数を、周波数によって高域の周波数スペクトル係数と低域の周波数スペクトル係数の2つのグループに分割する。そして、帯域分割部201は、低域の周波数スペクトル係数をIMDCT部202に供給し、高域の周波数スペクトル係数を無相関周波数時間変換部102に供給する。

[0137] IMDCT部202（第3の変換手段）は、帯域分割部201から供給される低域の周波数スペクトル係数に対してIMDCTを行い、低域の時間領域信号であるモノラル信号 $X_{M^{low}}$ （第3の時間領域信号）を得る。IMDCT部202は、低域のモノラル信号 $X_{M^{low}}$ を低域の左用のオーディオ信号として加算器203に供給するとともに、低域の右用のオーディオ信号として加算器204に供給する。

[0138] 加算器203には、帯域分割部201から出力された高域の周波数スペクトル係数に対して、無相関周波数時間変換部102およびステレオ合成部103による処理が行われた結果得られる高域の左用のオーディオ信号 X_L^{High} が入力される。加算器203は、その高域の左用のオーディオ信号 X_L^{High} と、IMDCT部202から低域の左用のオーディオ信号として供給される低域のモノラル信号 $X_{M^{low}}$ とを加算して、全周波数帯域の左用のオーディオ信号 X_L を生成する。

[0139] 加算器204には、帯域分割部201から出力された高域の周波数スペクトル係数に対して、無相関周波数時間変換部102およびステレオ合成部1

03による処理が行われた結果得られる高域の右用のオーディオ信号 X_{R}^{High} が入力される。加算器204は、その高域の右用のオーディオ信号 X_{R}^{High} と、IMDCT部202から低域の右用のオーディオ信号として供給される低域のモノラル信号 X_{M}^{Low} とを加算して、全周波数帯域の右用のオーディオ信号 X_R を出力する。

[0140] [音声処理装置の処理の説明]

図16は、図15の音声処理装置200による復号処理を説明するフローチャートである。この復号処理は、図2のオーディオ信号符号化部13を備える図1の符号化装置10と同様の空間符号化が行われ、高域についてのBCパラメータが多重化された符号化データが、音声処理装置200に入力されたとき、開始される。

[0141] 図16のステップS31乃至S33は、図14のステップS11乃至S13の処理と同様であるので、説明は繰り返しになるので省略する。

[0142] ステップS34において、帯域分割部201は、スペクトル逆量子化部53により得られる周波数スペクトル係数を、周波数によって高域の周波数スペクトル係数と低域の周波数スペクトル係数の2つのグループに分割する。そして、帯域分割部201は、低域の周波数スペクトル係数をIMDCT部202に供給し、高域の周波数スペクトル係数を無相関周波数時間変換部102に供給する。

[0143] ステップS35において、IMDCT部202は、帯域分割部201から供給される低域の周波数スペクトル係数に対してIMDCTを行い、低域の時間領域信号であるモノラル信号 X_{M}^{Low} を得る。IMDCT部202は、低域のモノラル信号 X_{M}^{Low} を低域の左用のオーディオ信号として加算器203に供給するとともに、低域の右用のオーディオ信号として加算器204に供給する。

[0144] ステップS36において、無相関周波数時間変換部102、ステレオ合成部103、および生成パラメータ計算部104は、帯域分割部201から供給される高域の周波数スペクトル係数に対してステレオ信号生成処理を行う。具体的には、無相関周波数時間変換部102、ステレオ合成部103、お

よび生成パラメータ計算部 104 は、図 14 のステップ S 14 乃至 S 18 の処理を行う。その結果得られる高域の左用のオーディオ信号 X_{L}^{High} は加算器 203 に入力され、高域の右用のオーディオ信号 X_{R}^{High} は加算器 204 に入力される。

[0145] ステップ S 37 において、加算器 203 は、IMDCT 部 202 から低域の左用のオーディオ信号として供給される低域のモノラル信号 X_M^{Low} と、無相関周波数時間変換部 102 から供給される高域の左用のオーディオ信号 X_{L}^{High} とを加算して、全周波数帯域の左用のオーディオ信号 X_L を生成する。そして、加算器 203 は、その全周波数帯域の左用のオーディオ信号 X_L を出力する。

[0146] ステップ S 38 において、加算器 204 は、IMDCT 部 202 から低域の右用のオーディオ信号として供給される低域のモノラル信号 X_M^{Low} と、無相関周波数時間変換部 102 から供給される高域の右用のオーディオ信号 X_{R}^{High} とを加算して、全周波数帯域の右用のオーディオ信号 X_R を生成する。そして、加算器 204 は、その全周波数帯域の右用のオーディオ信号 X_R を出力する。

[0147] 以上のように、音声処理装置 200 は、全周波数帯域のモノラル信号 X_M の符号化データを復号し、高域についてのみステレオ化する。これにより、低域のモノラル信号 X_M のステレオ化によって、音声が不自然になることを防止することができる。

[0148] なお、音声処理装置 200 では、帯域分割部 201 が、高域の周波数スペクトル係数と低域の周波数スペクトル係数に分割したが、所定の周波数帯域の周波数スペクトル係数と、それ以外の周波数帯域の周波数スペクトル係数に分割するようにしてもよい。即ち、ステレオ化の有無が、低域であるか、高域であるかによって選択されるのではなく、所定の周波数帯域であるか、それ以外の周波数帯域であるかによって選択されるようにしてもよい。

[0149] <第 3 実施の形態>

[音声処理装置の第 3 実施の形態の構成例]

図 17 は、本発明を適用した音声処理装置の第 3 実施の形態の構成例を示すブロック図である。

[0150] 図17に示す構成のうち、図4、図6、および図9の構成と同じ構成には同じ符号を付してある。重複する説明については適宜省略する。

[0151] 図17の音声処理装置300の構成は、主に、逆多重化部41と逆多重化部61の代わりに逆多重化部301が設けられている点、IMDCT部64-1乃至IMDCT部64-(N-1)の代わりにIMDCT部304-1乃至304-(N-1)が設けられている点、IMDCT部64-Nおよびステレオ信号生成部44の代わりにステレオ化部305が設けられている点、生成パラメータ計算部43、合成フィルタバンク65の代わりに生成パラメータ計算部104、合成フィルタバンク306が設けられている点が、図6のオーディオ信号復号部42と図7のステレオ信号生成部44を備える図4の復号装置40の構成と異なる。

[0152] 図17の音声処理装置300は、例えば図3のオーディオ信号符号化部13を備える図1の符号化装置10と同様の空間符号化が行われ、所定のサブバンド信号のBCパラメータが多重化された符号化データを復号する。

[0153] 具体的には、音声処理装置300の逆多重化部301は、図4の逆多重化部41と図6の逆多重化部61に対応する。即ち、逆多重化部301には、図3のオーディオ信号符号化部13を備える図1の符号化装置10と同様の空間符号化が行われ、所定のサブバンド信号のBCパラメータが多重化された符号化データが入力される。逆多重化部301は、入力された符号化データに対して逆多重化を行い、符号化データと所定のサブバンド信号のBCパラメータを得る。そして、逆多重化部301は、所定のサブバンド信号のBCパラメータを生成パラメータ計算部104に供給する。

[0154] また、逆多重化部301は、符号化データに対して逆多重化を行い、量子化され、エントロピー符号化されたN個のサブバンド信号の周波数スペクトル係数と量子化情報を得る。逆多重化部301は、量子化され、エントロピー符号化されたN個のサブバンド信号の周波数スペクトル係数をエントロピー復号部62に供給し、量子化情報をスペクトル逆量子化部63に供給する。

- [0155] IMDCT部304-1乃至304-(N-1) (第3の変換手段) およびステレオ化部305には、スペクトル逆量子化部63により復元されたN個のサブバンド信号の周波数スペクトル係数が、1つずつ入力される。
- [0156] IMDCT部304-1乃至304-(N-1)は、それぞれ、入力された周波数スペクトル係数に対してIMDCTを行い、周波数スペクトル係数を時間領域信号であるモノラル信号 X_M のサブバンド信号 X_{M^i} ($i=1, 2, \dots, N-1$)に変換する。IMDCT部304-1乃至304-(N-1)は、それぞれ、サブバンド信号 X_{M^i} を、左用のオーディオ信号 X_{L^i} と右用のオーディオ信号 X_{R^i} として合成フィルタバンク306に供給する。
- [0157] ステレオ化部305は、図9の無相関周波数時間変換部102とステレオ合成部103により構成される。ステレオ化部305は、生成パラメータ計算部104により生成された生成パラメータを用いて、スペクトル逆量子化部63から入力された所定のサブバンド信号の周波数スペクトル係数から、時間領域信号である左用のオーディオ信号のサブバンド信号 X_{L^A} と右用のオーディオ信号のサブバンド信号 X_{R^A} を生成する。そして、ステレオ化部305は、左用のサブバンド信号 X_{L^A} と右用のサブバンド信号 X_{R^A} を合成フィルタバンク306に供給する。
- [0158] 合成フィルタバンク306 (加算手段)は、左用のオーディオ信号のサブバンド信号を合成するための左用合成フィルタバンクと、右用のオーディオ信号のサブバンド信号を合成するための右用合成フィルタバンクにより構成される。合成フィルタバンク306の左用合成フィルタバンクは、IMDCT部304-1乃至304-(N-1)からの左用のサブバンド信号 X_{L^1} 乃至 $X_{L^{N-1}}$ と、ステレオ化部305からの左用のサブバンド信号 X_{L^A} を合成する。そして、左用合成フィルタバンクは、合成の結果得られる全周波数帯域の左用のオーディオ信号 X_L を出力する。
- [0159] また、合成フィルタバンク306の右用合成フィルタバンクは、IMDCT部304-1乃至304-(N-1)からの右用のサブバンド信号 X_{R^1} 乃至 $X_{R^{N-1}}$ と、ステレオ化部305からの右用のサブバンド信号 X_{R^A} を合成する。そし

て、右用合成フィルタバンクは、合成の結果得られる全周波数帯域の右用のオーディオ信号 X_R を出力する。

[0160] なお、図17の音声処理装置300では、1つのサブバンド信号についてのみステレオ化が行われるようにしたが、複数のサブバンド信号についてステレオ化が行われるようにすることもできる。また、ステレオ化が行われるサブバンド信号は、予め設定されるのではなく、符号化側で動的に設定されるようにしてもよい。この場合、例えば、BCパラメータにステレオ化の対象となるサブバンド信号を特定する情報が含まれる。

[0161] [音声処理装置の処理の説明]

図18は、図17の音声処理装置300による復号処理を説明するフローチャートである。この復号処理は、例えば、図3のオーディオ信号符号化部13を備える図1の符号化装置10と同様の空間符号化が行われ、所定のサブバンド信号のBCパラメータが多重化された符号化データが音声処理装置300に入力されたとき、開始される。

[0162] 図18のステップS51において、逆多重化部301は、入力された多重化された符号化データに対して逆多重化を行い、符号化データと所定のサブバンド信号のBCパラメータを得る。そして、逆多重化部301は、所定のサブバンド信号のBCパラメータを生成パラメータ計算部104に供給する。また、逆多重化部301は、符号化データに対して逆多重化を行い、量子化され、エントロピー符号化されたN個のサブバンド信号の周波数スペクトル係数と量子化情報を得る。逆多重化部301は、量子化され、エントロピー符号化されたN個のサブバンド信号の周波数スペクトル係数をエントロピー復号部62に供給し、量子化情報をスペクトル逆量子化部63に供給する。

[0163] ステップS52において、エントロピー復号部62は、逆多重化部101から供給されるN個のサブバンド信号の周波数スペクトル係数に対してエントロピー復号を行い、スペクトル逆量子化部63に供給する。

[0164] ステップS53において、スペクトル逆量子化部63は、逆多重化部30

1から供給される量子化情報に基づいて、エントロピー復号部62から供給されるエントロピー復号の結果得られたN個のサブバンド信号の周波数スペクトル係数それぞれに対して逆量子化を行う。そして、スペクトル逆量子化部63は、その結果復元されたN個のサブバンド信号の周波数スペクトル係数を、IMDCT部304-1乃至304-(N-1)およびステレオ化部305に1つずつ供給する。

[0165] ステップS54において、IMDCT部304-1乃至304-(N-1)は、それぞれ、スペクトル逆量子化部63から供給される周波数スペクトル係数に対してIMDCTを行う。そして、IMDCT部304-1乃至304-(N-1)は、それぞれ、その結果得られるモノラル信号のサブバンド信号 X_M^i ($i=1, 2, \dots, N-1$)を、左用のオーディオ信号のサブバンド信号 X_L^i と右用のオーディオ信号のサブバンド信号 X_R^i として合成フィルタバンク306に供給する。

[0166] ステップS55において、ステレオ化部305は、生成パラメータ計算部104から供給される生成パラメータを用いて、スペクトル逆量子化部63から供給される所定のサブバンド信号の周波数スペクトル係数に対して、ステレオ信号生成処理を行う。そして、ステレオ化部305は、その結果得られる時間領域信号である左用のオーディオ信号のサブバンド信号 X_L^A と右用のオーディオ信号のサブバンド信号 X_R^A を合成フィルタバンク306に供給する。

[0167] ステップS56において、合成フィルタバンク306の左用合成フィルタバンクは、IMDCT部304-1乃至304-(N-1)とステレオ化部305からそれぞれ供給される左用のオーディオ信号の全サブバンド信号を合成して、全周波数帯域の左用のオーディオ信号 X_L を生成する。そして、左用合成フィルタバンクは、その全周波数帯域の左用のオーディオ信号 X_L を出力する。

[0168] ステップS57において、合成フィルタバンク306の右用合成フィルタバンクは、IMDCT部304-1乃至304-(N-1)とステレオ化部305

からそれぞれ供給される右用のオーディオ信号の全サブバンド信号を合成して、全周波数帯域の右用のオーディオ信号 X_R を生成する。そして、右用合成フィルタバンクは、その全周波数帯域の右用のオーディオ信号 X_R を出力する。

[0169] <第4実施の形態>

[音声処理装置の第4実施の形態の構成例]

図19は、本発明を適用した音声処理装置の第4実施の形態の構成例を示すブロック図である。

[0170] 図19に示す構成のうち、図15の構成と同じ構成には同じ符号を付してある。重複する説明については適宜省略する。

[0171] 図19の音声処理装置400の構成は、主に、帯域分割部201の代わりにスペクトル分離部401が設けられ、IMDCT部202の代わりにIMDCT402および403が設けられ、加算器203、加算器204の代わりに加算器404、加算器405が設けられている点が、図15の構成と異なる。

[0172] 音声処理装置400は、インテンシティ符号化された符号化データであって、従来のチャンネル間の周波数スペクトル係数のレベル比の代わりにインテンシティ開始周波数 F_{is} 以上の周波数のBCパラメータが多重化された符号化データを復号する。

[0173] 即ち、音声処理装置400によって復号される符号化データは、例えば、符号化対象のステレオ信号をモノラル信号 X_M にダウンミックスし、その結果得られるモノラル信号 X_M と符号化対象のステレオ信号のインテンシティ開始周波数 F_{is} 以上の周波数の成分をハイパスフィルタ等によって抽出してBCパラメータを検出する符号化装置により生成される。

[0174] 音声処理装置400のスペクトル分離部401（分離手段）は、スペクトル逆量子化部53により復元された周波数スペクトル係数を得る。スペクトル分離部401は、その周波数スペクトル係数を、インテンシティ開始周波数 F_{is} より低い周波数のステレオ信号の周波数スペクトル係数と、インテンシティ開始周波数 F_{is} 以上の周波数のモノラル信号 $X_{M^{high}}$ の周波数スペクトル

係数に分離する。スペクトル分離部 401 は、インテンシティ開始周波数 F_{is} より低い周波数のステレオ信号の左用のオーディオ信号 $X_{L^{low}}$ の周波数スペクトル係数を IMDCT 部 402 に供給し、右用のオーディオ信号 $X_{R^{low}}$ の周波数スペクトル係数を IMDCT 部 403 に供給する。また、スペクトル分離部 401 は、モノラル信号 $X_{M^{high}}$ の周波数スペクトル係数を無相関周波数時間変換部 102 に供給する。

[0175] IMDCT 部 402 (第 3 の変換手段) は、スペクトル分離部 401 から供給される左用のオーディオ信号 $X_{L^{low}}$ の周波数スペクトル係数に対して IMDCT を行い、その結果得られる左用のオーディオ信号 $X_{L^{low}}$ を加算器 404 に供給する。

[0176] IMDCT 部 403 (第 3 の変換手段) は、スペクトル分離部 401 から供給される右用のオーディオ信号 $X_{R^{low}}$ の周波数スペクトル係数に対して IMDCT を行い、その結果得られる右用のオーディオ信号 $X_{R^{low}}$ を加算器 405 に供給する。

[0177] 加算器 404 (加算手段) は、ステレオ合成部 103 により生成されるインテンシティ開始周波数 F_{is} 以上の周波数の時間領域信号である左用のオーディオ信号 $X_{L^{high}}$ と、IMDCT 部 402 から供給される左用のオーディオ信号 $X_{L^{low}}$ とを加算する。加算器 404 は、その結果得られるオーディオ信号を全周波数帯域の左用のオーディオ信号 X_L として出力する。

[0178] 加算器 405 (加算手段) は、ステレオ合成部 103 により生成されるインテンシティ開始周波数 F_{is} 以上の周波数の時間領域信号である右用のオーディオ信号 $X_{R^{high}}$ と、IMDCT 部 402 から供給される右用のオーディオ信号 $X_{R^{low}}$ とを加算する。加算器 405 は、その結果得られるオーディオ信号を全周波数帯域の右用のオーディオ信号 X_R として出力する。

[0179] 以上のように、音声処理装置 400 は、インテンシティ符号化された符号化データに多重化された BC パラメータを用いて、インテンシティ符号化によってモノラル化されたインテンシティ開始周波数 F_{is} 以上の周波数の成分をステレオ化する。これにより、従来のチャンネル間の周波数スペクトル係数

のレベル比を用いてステレオ化を行うインテンシティ復号装置に比べて、インテンシティ開始周波数 F_{is} 以上の周波数の成分のステレオ感を復元することができる。

[0180] [音声処理装置の処理の説明]

図20は、図19の音声処理装置400による復号処理を説明するフローチャートである。この復号処理は、例えば、インテンシティ符号化され、インテンシティ開始周波数 F_{is} 以上の周波数のBCパラメータが多重化された符号化データが入力されたとき、開始される。

[0181] 図20のステップS71乃至S73の処理は、図16のステップS31乃至S33の処理と同様であるので、説明は省略する。

[0182] ステップS74において、スペクトル分離部401は、スペクトル逆量子化部53により復元された周波数スペクトル係数を、インテンシティ開始周波数 F_{is} より低い周波数のステレオ信号の周波数スペクトル係数と、インテンシティ開始周波数 F_{is} 以上の周波数のモノラル信号 $X_{M^{high}}$ の周波数スペクトル係数に分離する。スペクトル分離部401は、インテンシティ開始周波数 F_{is} より低い周波数のステレオ信号の左用のオーディオ信号 $X_{L^{low}}$ の周波数スペクトル係数をIMDCT部402に供給し、右用のオーディオ信号 $X_{R^{low}}$ の周波数スペクトル係数をIMDCT部403に供給する。また、スペクトル分離部401は、モノラル信号 $X_{M^{high}}$ の周波数スペクトル係数を無相関周波数時間変換部102に供給する。

[0183] ステップS75において、IMDCT部402は、スペクトル分離部401から供給される左用のオーディオ信号 $X_{L^{low}}$ の周波数スペクトル係数に対してIMDCTを行う。そして、IMDCT部402は、その結果得られる左用のオーディオ信号 $X_{L^{low}}$ を加算器404に供給する。

[0184] ステップS76において、IMDCT部403は、スペクトル分離部401から供給される右用のオーディオ信号 $X_{R^{low}}$ の周波数スペクトル係数に対してIMDCTを行う。そして、IMDCT部403は、その結果得られる右用のオーディオ信号 $X_{R^{low}}$ を加算器405に供給する。

- [0185] ステップS 7 7において、無相関周波数時間変換部 1 0 2、ステレオ合成部 1 0 3、および生成パラメータ計算部 1 0 4は、スペクトル分離部 4 0 1からのモノラル信号 $X_{M^{high}}$ の周波数スペクトル係数に対してステレオ信号生成処理を行う。その結果得られる時間領域信号である左用のオーディオ信号 $X_{L^{high}}$ は、加算器 4 0 4に供給され、右用のオーディオ信号 $X_{R^{high}}$ は、加算器 4 0 5に供給される。
- [0186] ステップS 7 8において、加算器 4 0 4は、IMDCT部 4 0 2からのインテンシティ開始周波数 F_{is} より低い周波数の左用のオーディオ信号 $X_{L^{low}}$ とステレオ合成部 1 0 3からのインテンシティ開始周波数 F_{is} 以上の周波数の左用のオーディオ信号 $X_{L^{high}}$ とを加算して、全周波数帯域の左用のオーディオ信号 X_L を生成する。そして、加算器 4 0 4は、その左用のオーディオ信号 X_L を出力する。
- [0187] ステップS 7 9において、加算器 4 0 5は、IMDCT部 4 0 3からのインテンシティ開始周波数 F_{is} より低い周波数の右用のオーディオ信号 $X_{R^{low}}$ と、ステレオ合成部 1 0 3からのインテンシティ開始周波数 F_{is} 以上の周波数の右用のオーディオ信号 $X_{R^{high}}$ とを加算して、全周波数帯域の右用のオーディオ信号 X_R を生成する。そして、加算器 4 0 5は、その右用のオーディオ信号 X_R を出力する。
- [0188] なお、上述した説明では、音声処理装置 1 0 0（2 0 0， 3 0 0， 4 0 0）が、MDCTによって時間周波数変換された符号化データを復号するようにしたので、周波数時間変換時にIMDCTが行われたが、MDSTによって時間周波数変換された符号化データを復号する場合には、周波数時間変換時にIMDSTが行われる。
- [0189] また、上述した説明では、無相関時間周波数変換部 1 0 2において、基底が互いに直交する変換としてIMDCT変換とIMDST変換が用いられたが、サイン変換とコサイン変換等の他の重複直交変換が用いられてもよい。
- [0190] [本発明を適用したコンピュータの説明]
- 次に、上述した一連の処理は、ハードウェアにより行うこともできるし、

ソフトウェアにより行うこともできる。一連の処理をソフトウェアによって行う場合には、そのソフトウェアを構成するプログラムが、汎用のコンピュータ等にインストールされる。

[0191] そこで、図 2 1 は、上述した一連の処理を実行するプログラムがインストールされるコンピュータの一実施の形態の構成例を示している。

[0192] プログラムは、コンピュータに内蔵されている記録媒体としての記憶部 5 0 8 や ROM (Read Only Memory) 5 0 2 に予め記録しておくことができる。

[0193] あるいはまた、プログラムは、リムーバブルメディア 5 1 1 に格納（記録）しておくことができる。このようなリムーバブルメディア 5 1 1 は、いわゆるパッケージソフトウェアとして提供することができる。ここで、リムーバブルメディア 5 1 1 としては、例えば、フレキシブルディスク、CD-ROM (Compact Disc Read Only Memory)、MO (Magneto Optical) ディスク、DVD (Digital Versatile Disc)、磁気ディスク、半導体メモリ等がある。

[0194] なお、プログラムは、上述したようなリムーバブルメディア 5 1 1 からドライブ 5 1 0 を介してコンピュータにインストールする他、通信網や放送網を介して、コンピュータにダウンロードし、内蔵する記憶部 5 0 8 にインストールすることができる。すなわち、プログラムは、例えば、ダウンロードサイトから、デジタル衛星放送用の人工衛星を介して、コンピュータに無線で転送したり、LAN (Local Area Network)、インターネットといったネットワークを介して、コンピュータに有線で転送することができる。

[0195] コンピュータは、CPU (Central Processing Unit) 5 0 1 を内蔵しており、CPU 5 0 1 には、バス 5 0 4 を介して、入出力インタフェース 5 0 5 が接続されている。

[0196] CPU 5 0 1 は、入出力インタフェース 5 0 5 を介して、ユーザによって、入力部 5 0 6 が操作等されることにより指令が入力されると、それに従って、ROM 5 0 2 に格納されているプログラムを実行する。あるいは、CPU 5 0 1 は、記憶部 5 0 8 に格納されたプログラムを、RAM (Random Access Memory) 5 0 3 にロードして実行する。

[0197] これにより、CPU 501は、上述したフローチャートにしたがった処理、あるいは上述したブロック図の構成により行われる処理を行う。そして、CPU 501は、その処理結果を、必要に応じて、例えば、入出力インタフェース 505を介して、出力部 507から出力、あるいは、通信部 509から送信、さらには、記憶部 508に記録等させる。

[0198] なお、入力部 506は、キーボードや、マウス、マイク等で構成される。また、出力部 507は、LCD(Liquid Crystal Display)やスピーカ等で構成される。

[0199] ここで、本明細書において、コンピュータがプログラムに従って行う処理は、必ずしもフローチャートとして記載された順序に沿って時系列に行われる必要はない。すなわち、コンピュータがプログラムに従って行う処理は、並列的あるいは個別に実行される処理（例えば、並列処理あるいはオブジェクトによる処理）も含む。

[0200] また、プログラムは、1のコンピュータ（プロセッサ）により処理されるものであっても良いし、複数のコンピュータによって分散処理されるものであっても良い。さらに、プログラムは、遠方のコンピュータに転送されて実行されるものであっても良い。

[0201] 本発明は、オーディオ信号の擬似ステレオ化技術に適用することができる。

[0202] 本発明の実施の形態は、上述した実施の形態に限定されるものではなく、本発明の要旨を逸脱しない範囲において種々の変更が可能である。

符号の説明

[0203] 54 IMDCT部, 100 音声処理装置, 101 逆多重化部, 103 ステレオ合成部, 111 IMDST部, 121 スペクトル反転部, 122 IMDCT部, 123 符号反転部, 200 音声処理装置, 201 帯域分割部, 202 IMDCT部, 203, 204 加算器, 300 音声処理装置, 301 逆多重化部, 304-1乃至304-N IMDCT部, 305 ステレオ化部, 306 合成フィルタバンク, 400

音声処理装置, 401 スペクトル分離部, 402, 403 IMDCT部
, 404, 405 加算器

請求の範囲

[請求項1]

複数チャンネルの音声の時間領域信号である音声信号から生成された前記複数チャンネルより少ないチャンネルの音声信号の周波数領域の係数と、前記複数チャンネルのチャンネル間の関係を表すパラメータとを取得する取得手段と、

前記取得手段により取得された前記周波数領域の係数を、第1の時間領域信号に変換する第1の変換手段と、

前記取得手段により取得された前記周波数領域の係数を、第2の時間領域信号に変換する第2の変換手段と、

前記パラメータを用いて前記第1の時間領域信号と前記第2の時間領域信号を合成することにより、前記複数チャンネルの音声信号を生成する合成手段と

を備え、

前記第1の変換手段による変換における基底と前記第2の変換手段による変換における基底は直交する

音声処理装置。

[請求項2]

前記取得手段により取得された前記周波数領域の係数を、周波数によって複数のグループに分割する分割手段と、

前記複数のグループのうちの第1のグループに分割された前記周波数領域の係数を、第3の時間領域信号に変換する第3の変換手段と、

前記第3の時間領域信号を前記第1のグループの周波数帯域の各チャンネルの音声信号とし、その第3の時間領域信号と、前記合成手段により生成された前記複数チャンネルの音声信号とをチャンネルごとに加算して、全周波数帯域の前記複数チャンネルの音声信号を生成する加算手段と

をさらに備え、

前記取得手段は、前記周波数領域の係数と前記第1のグループ以外のグループである第2のグループの周波数帯域の前記パラメータを取

得し、

前記第 1 の変換手段は、前記第 2 のグループに分割された前記周波数領域の係数を、前記第 1 の時間領域信号に変換し、

前記第 2 の変換手段は、前記第 2 のグループに分割された前記周波数領域の係数を、前記第 2 の時間領域信号に変換し、

前記合成手段は、前記パラメータを用いて前記第 1 の時間領域信号と前記第 2 の時間領域信号を合成することにより、前記第 2 のグループの周波数帯域の前記複数チャンネルの音声信号を生成する

請求項 1 に記載の音声処理装置。

[請求項 3]

前記取得手段により取得される周波数により複数のグループに分割された前記周波数領域の係数のうちの第 1 のグループの周波数領域の係数を、第 3 の時間領域信号に変換する第 3 の変換手段と、

前記第 3 の時間領域信号を前記第 1 のグループの周波数帯域の各チャンネルの音声信号とし、その第 3 の時間領域信号と、前記合成手段により生成された前記複数チャンネルの音声信号とをチャンネルごとに加算して、全周波数帯域の前記複数チャンネルの音声信号を生成する加算手段と

をさらに備え、

前記取得手段は、各グループの前記周波数領域の係数と、前記複数のグループのうちの前記第 1 のグループ以外のグループである第 2 のグループの周波数帯域の前記パラメータを取得し、

前記第 1 の変換手段は、前記第 2 のグループに分割された前記周波数領域の係数を、前記第 1 の時間領域信号に変換し、

前記第 2 の変換手段は、前記第 2 のグループに分割された前記周波数領域の係数を、前記第 2 の時間領域信号に変換し、

前記合成手段は、前記パラメータを用いて前記第 1 の時間領域信号と前記第 2 の時間領域信号を合成することにより、前記第 2 のグループの周波数帯域の前記複数チャンネルの音声信号を生成する

請求項 1 に記載の音声処理装置。

[請求項4] 前記周波数領域の係数は、前記複数チャンネルの音声信号の周波数領域の係数から生成される

請求項 1 に記載の音声処理装置。

[請求項5] 前記取得手段により取得された所定の周波数帯域の前記周波数領域の係数と、その周波数帯域以外の周波数帯域の前記複数チャンネルの音声信号の周波数領域の係数を分離する分離手段と、

前記分離手段により分離された前記複数チャンネルの音声信号の周波数領域の係数を、前記複数チャンネルの第 3 の時間領域信号に変換する第 3 の変換手段と、

前記複数チャンネルの第 3 の時間領域信号を前記所定の周波数帯域以外の周波数帯域の前記複数チャンネルの音声信号とし、その第 3 の時間領域信号と、前記合成手段により生成された前記複数チャンネルの音声信号とをチャンネルごとに加算して、全周波数帯域の前記複数チャンネルの音声信号を生成する加算手段と

をさらに備え、

前記取得手段は、前記所定の周波数帯域の前記周波数領域の係数、前記所定の周波数帯域以外の周波数帯域の前記複数チャンネルの音声信号の周波数領域の係数、および、前記所定の周波数帯域の前記パラメータを取得し、

前記第 1 の変換手段は、前記分離手段により分離された前記所定の周波数帯域の前記周波数領域の係数を、前記第 1 の時間領域信号に変換し、

前記第 2 の変換手段は、前記分離手段により分離された前記所定の周波数帯域の前記周波数領域の係数を、前記第 2 の時間領域信号に変換し、

前記合成手段は、前記パラメータを用いて前記第 1 の時間領域信号と前記第 2 の時間領域信号を合成することにより、前記所定の周波数

帯域の前記複数チャンネルの音声信号を生成する

請求項 4 に記載の音声処理装置。

[請求項6] 前記周波数領域の係数は、MDCT (Modified Discrete Cosine Transform) 係数であり、

前記第 1 の変換手段による変換は、IMDCT (Inverse Modified Discrete Cosine Transform) であり、

前記第 2 の変換手段による変換は、IMDST (Inverse Modified Discrete Sine Transform) である

請求項 1 乃至 5 のいずれかに記載の音声処理装置。

[請求項7] 前記第 2 の変換手段は、

前記周波数領域の係数を周波数が逆順になるように反転するスペクトル反転手段と

前記スペクトル反転手段による反転の結果得られる周波数領域の係数にIMDCT (Inverse Modified Discrete Cosine Transform) を行い、時間領域信号を得るIMDCT手段と、

前記IMDCT手段により得られた時間領域信号の各サンプルの符号を1つ置きに反転する符号反転手段と

を備え、

前記周波数領域の係数は、MDCT (Modified Discrete Cosine Transform) 係数であり、

前記第 1 の変換手段による変換は、IMDCTである

請求項 1 乃至 5 のいずれかに記載の音声処理装置。

[請求項8] 音声処理装置が、

複数チャンネルの音声の時間領域信号である音声信号から生成された前記複数チャンネルより少ないチャンネルの音声信号の周波数領域の係数と、前記複数チャンネルのチャンネル間の関係を表すパラメータとを取得する取得ステップと、

前記取得ステップの処理により取得された前記周波数領域の係数を

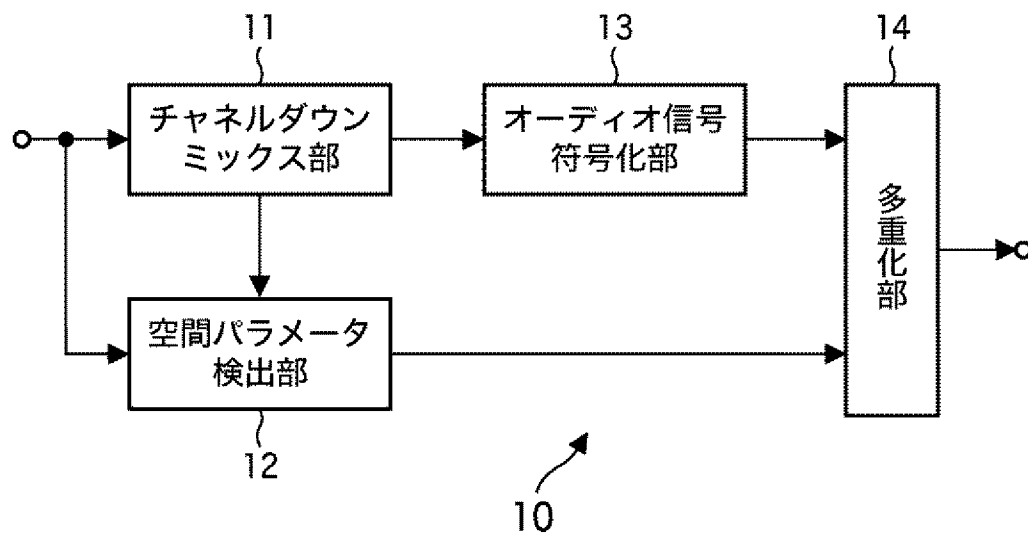
、第 1 の時間領域信号に変換する第 1 の変換ステップと、
前記取得ステップの処理により取得された前記周波数領域の係数を
、第 2 の時間領域信号に変換する第 2 の変換ステップと、
前記パラメータを用いて前記第 1 の時間領域信号と前記第 2 の時間
領域信号を合成することにより、前記複数チャンネルの音声信号を生
成する合成ステップと
を含み、
前記第 1 の変換ステップの処理による変換における基底と前記第 2
の変換ステップの処理による変換における基底は直交する
音声信号処理方法。

[請求項9]

コンピュータに、
複数チャンネルの音声の時間領域信号である音声信号から生成され
た前記複数チャンネルより少ないチャンネルの音声信号の周波数領域
の係数と、前記複数チャンネルのチャンネル間の関係を表すパラメー
タとを取得する取得ステップと、
前記取得ステップの処理により取得された前記周波数領域の係数を
、第 1 の時間領域信号に変換する第 1 の変換ステップと、
前記取得ステップの処理により取得された前記周波数領域の係数を
、第 2 の時間領域信号に変換する第 2 の変換ステップと、
前記パラメータを用いて前記第 1 の時間領域信号と前記第 2 の時間
領域信号を合成することにより、前記複数チャンネルの音声信号を生
成する合成ステップと
を含み、
前記第 1 の変換ステップの処理による変換における基底と前記第 2
の変換ステップの処理による変換における基底は直交する
処理を実行させるためのプログラム。

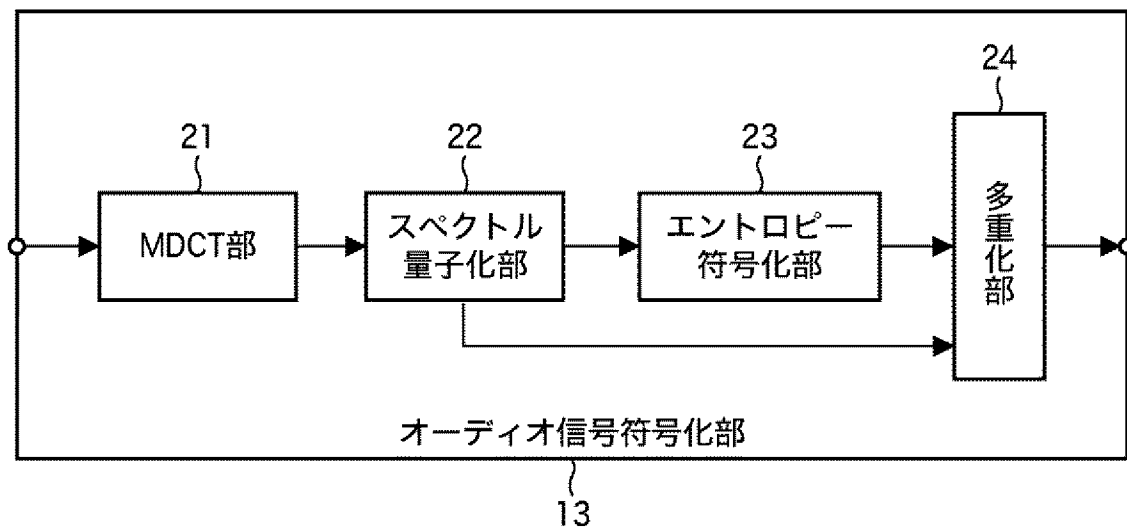
[図1]

図 1



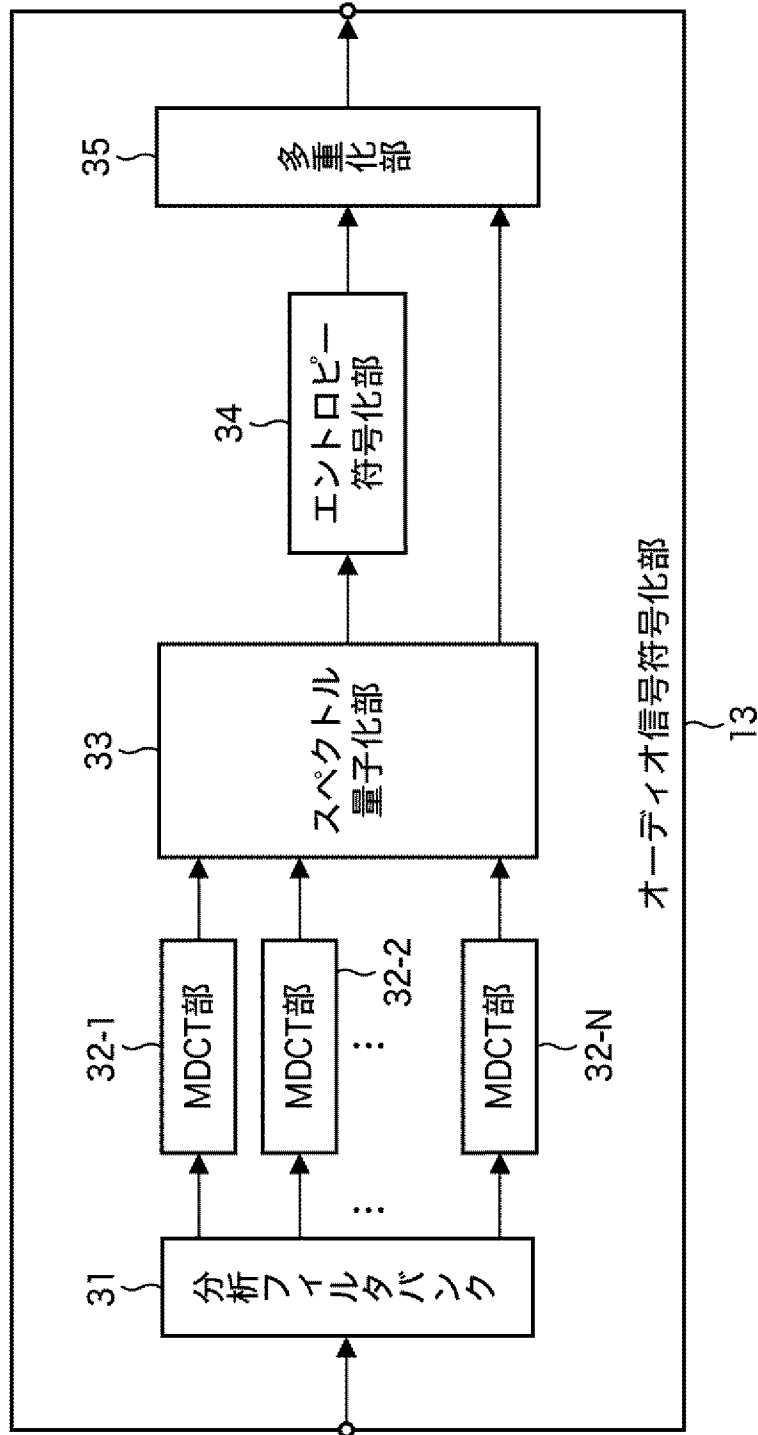
[図2]

図 2



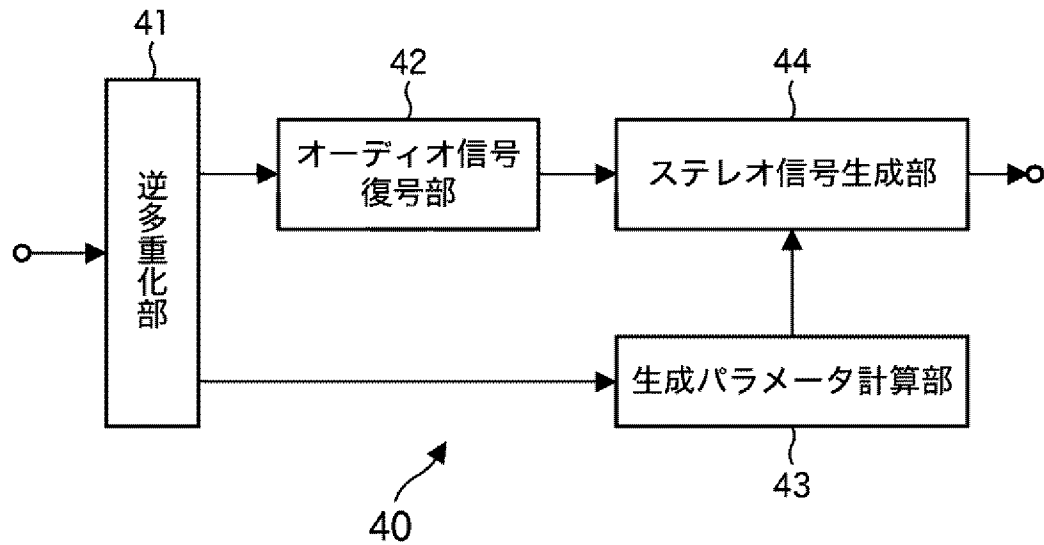
[図3]

図3



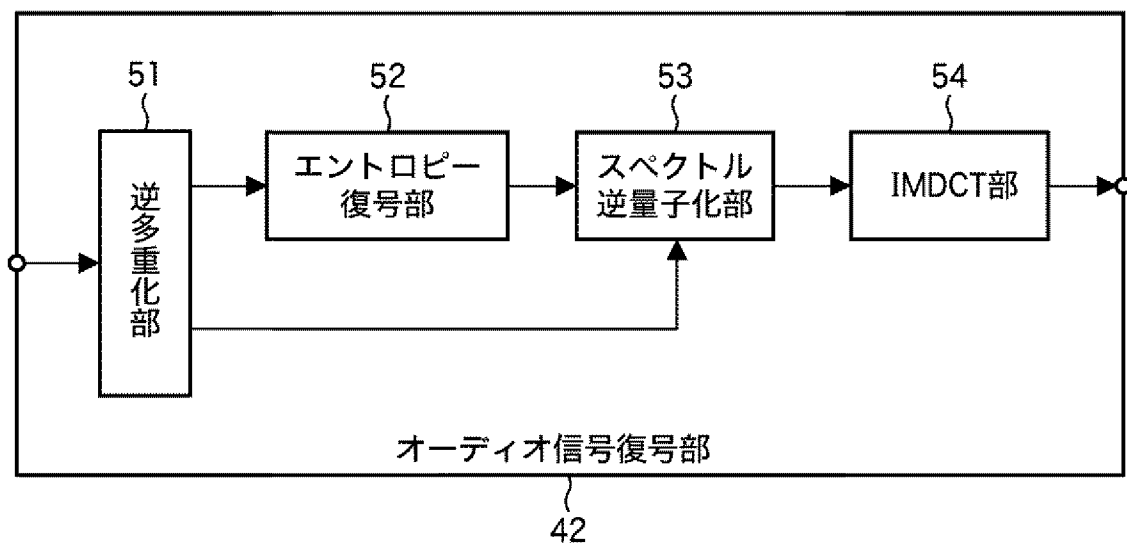
[図4]

図 4



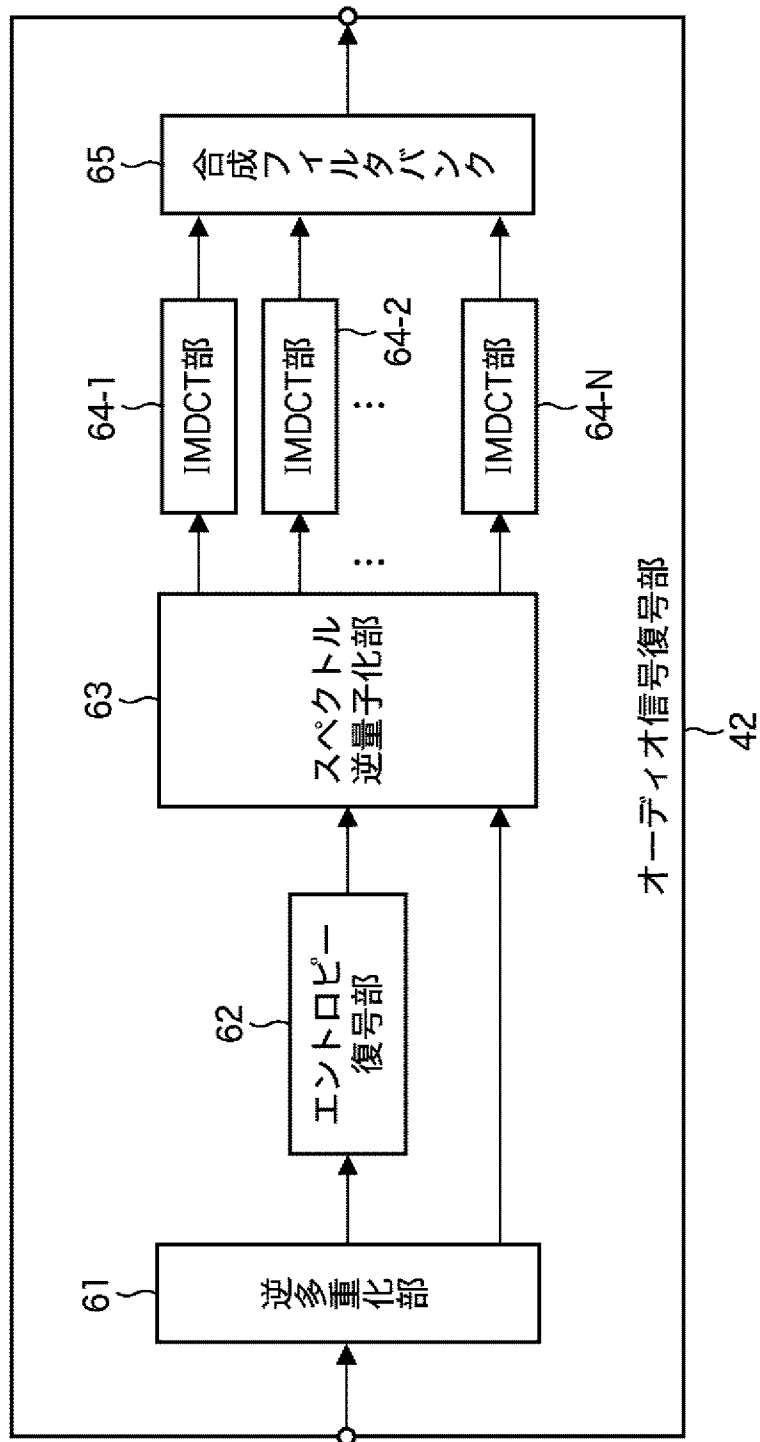
[図5]

図 5



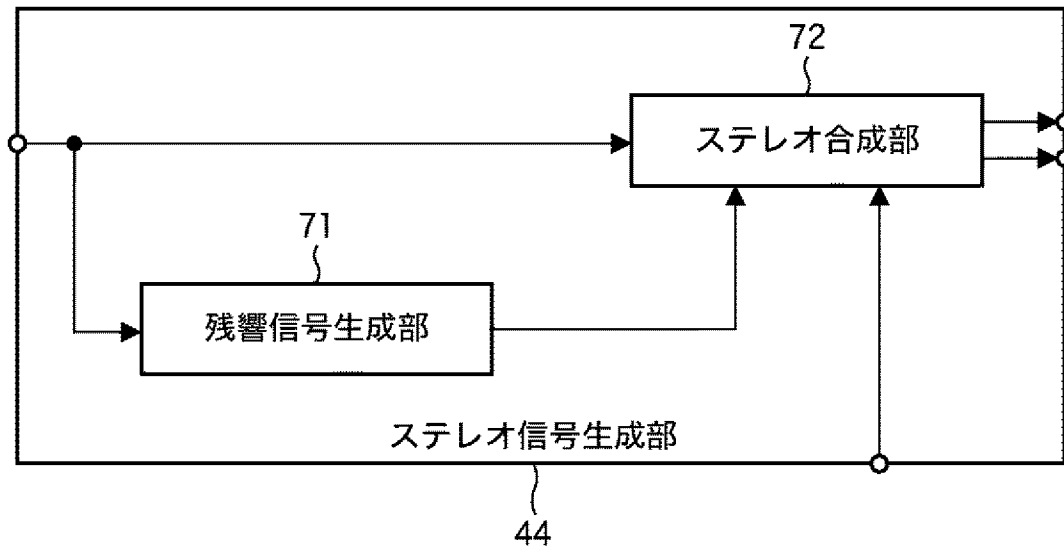
[図6]

図 6



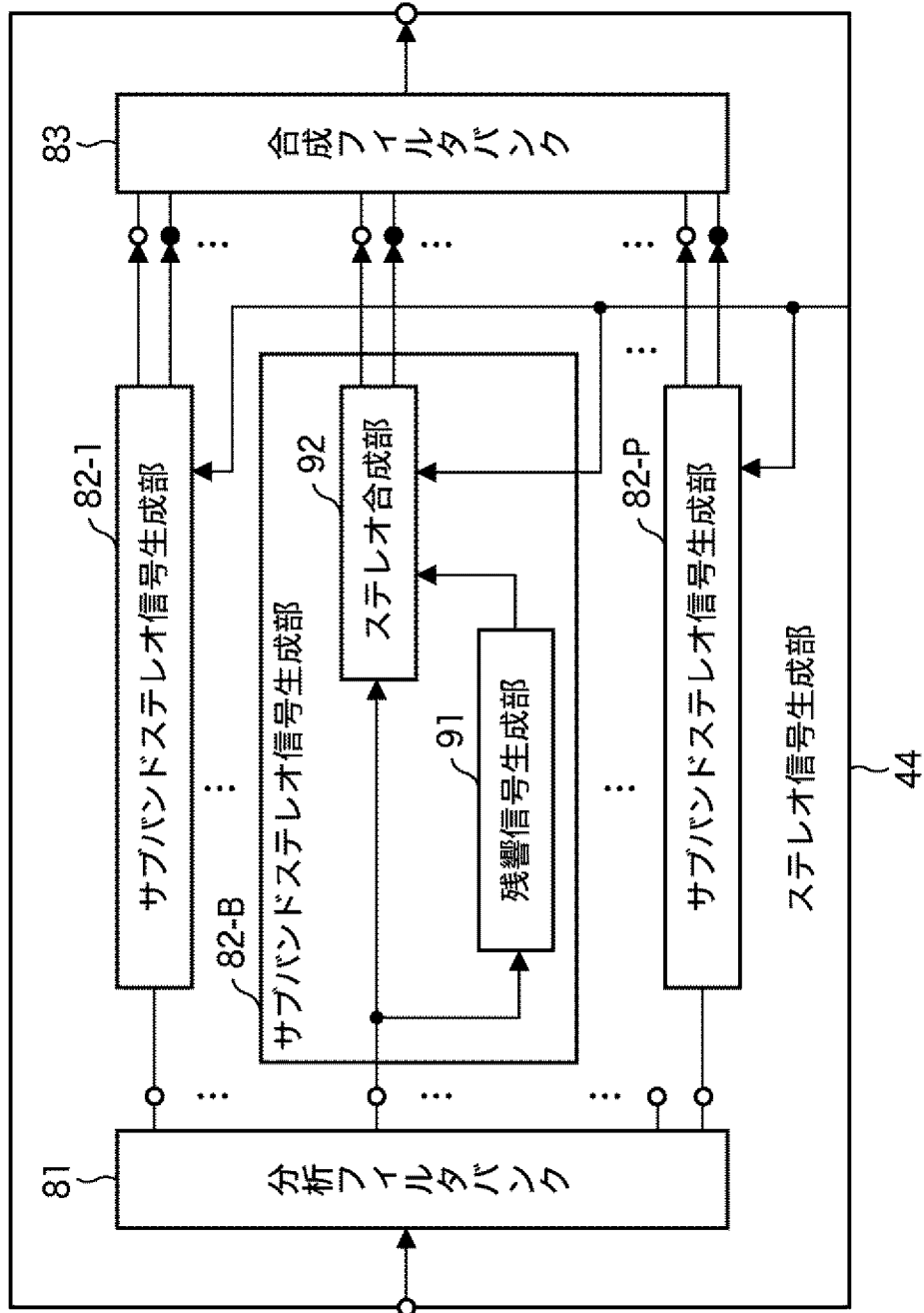
[図7]

図 7



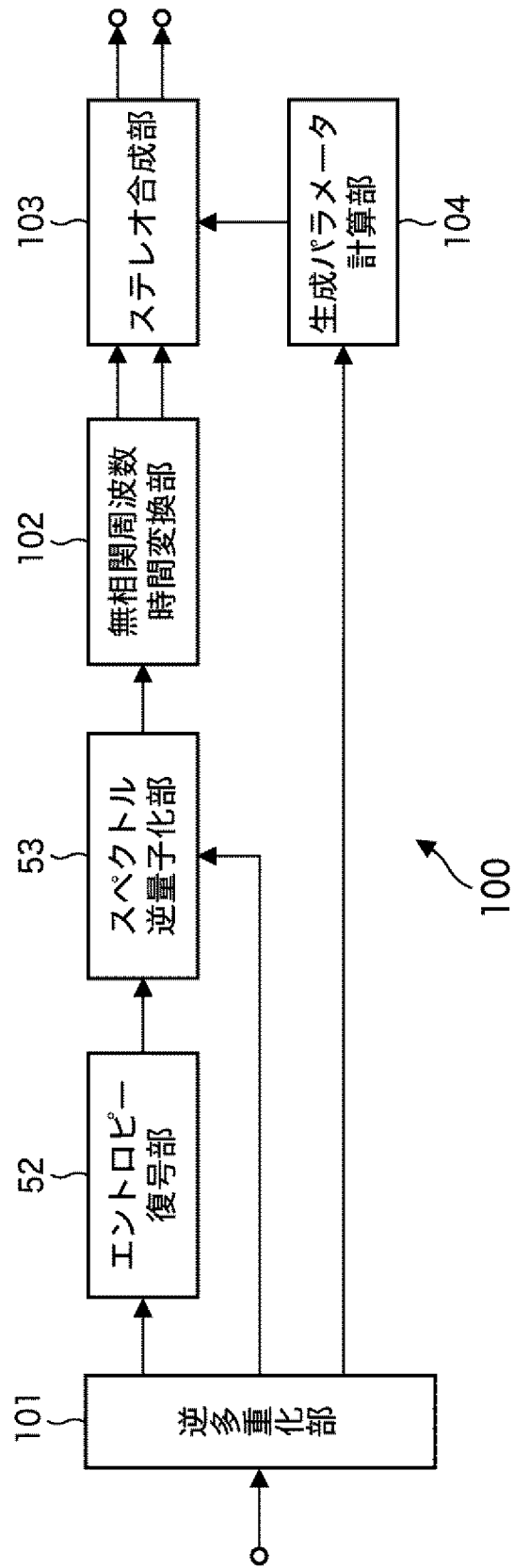
[図8]

図 8



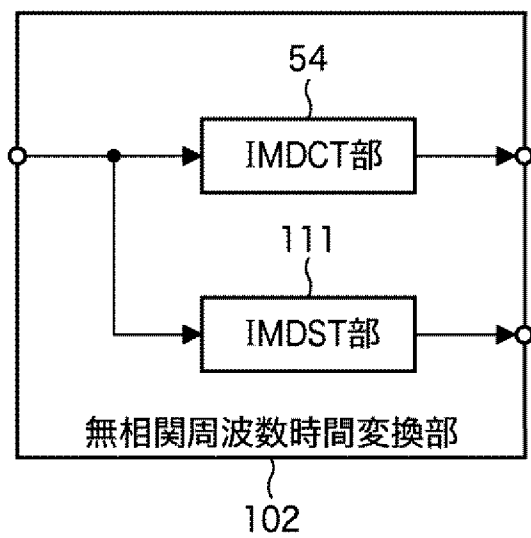
[図9]

図9



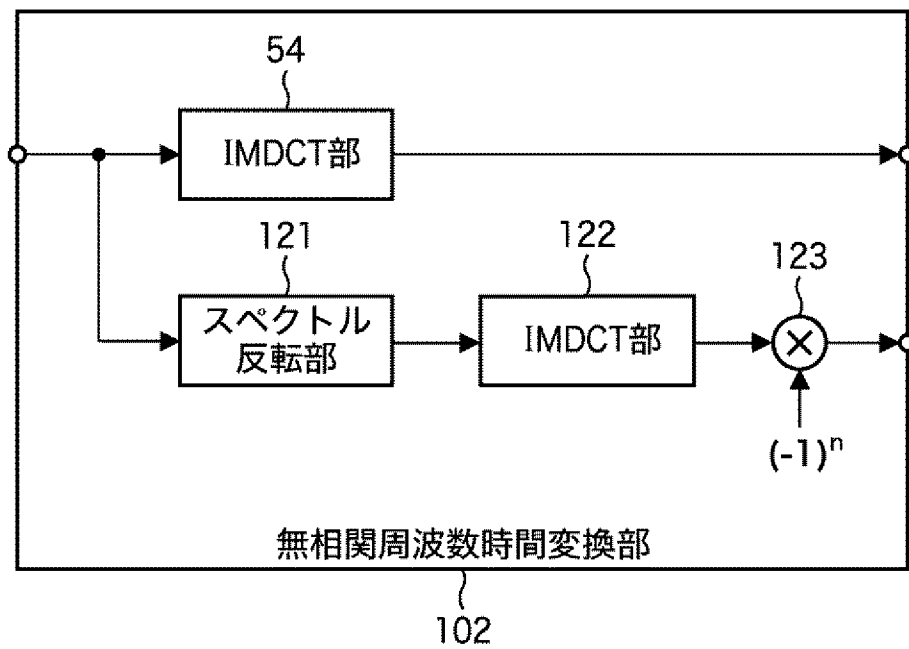
[図10]

図 10



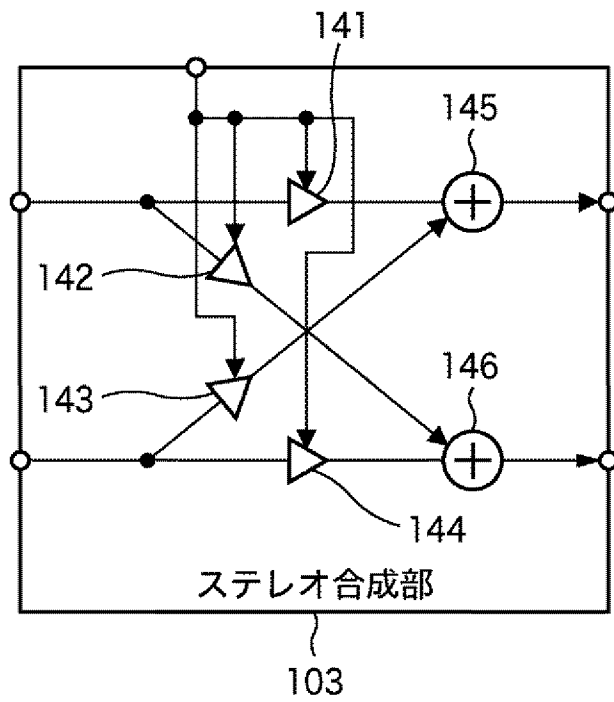
[図11]

図 11



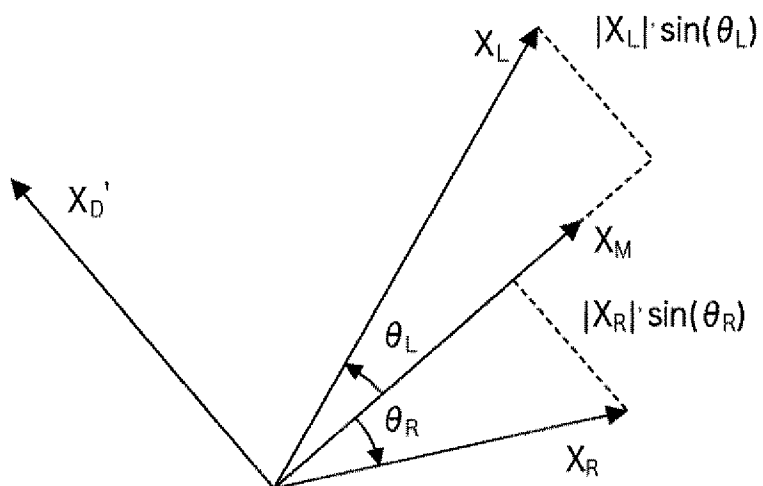
[図12]

図 12

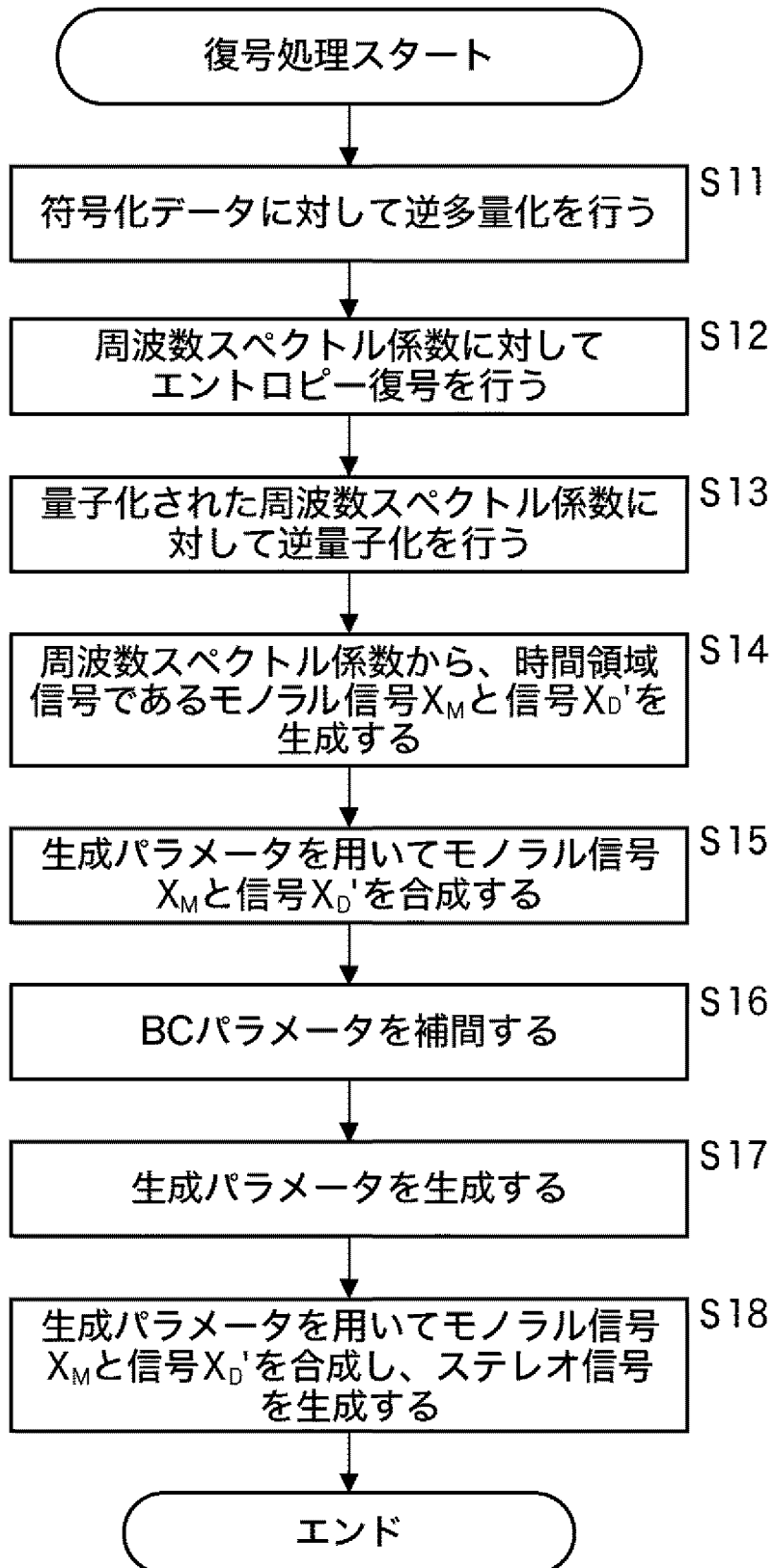


[図13]

図 13

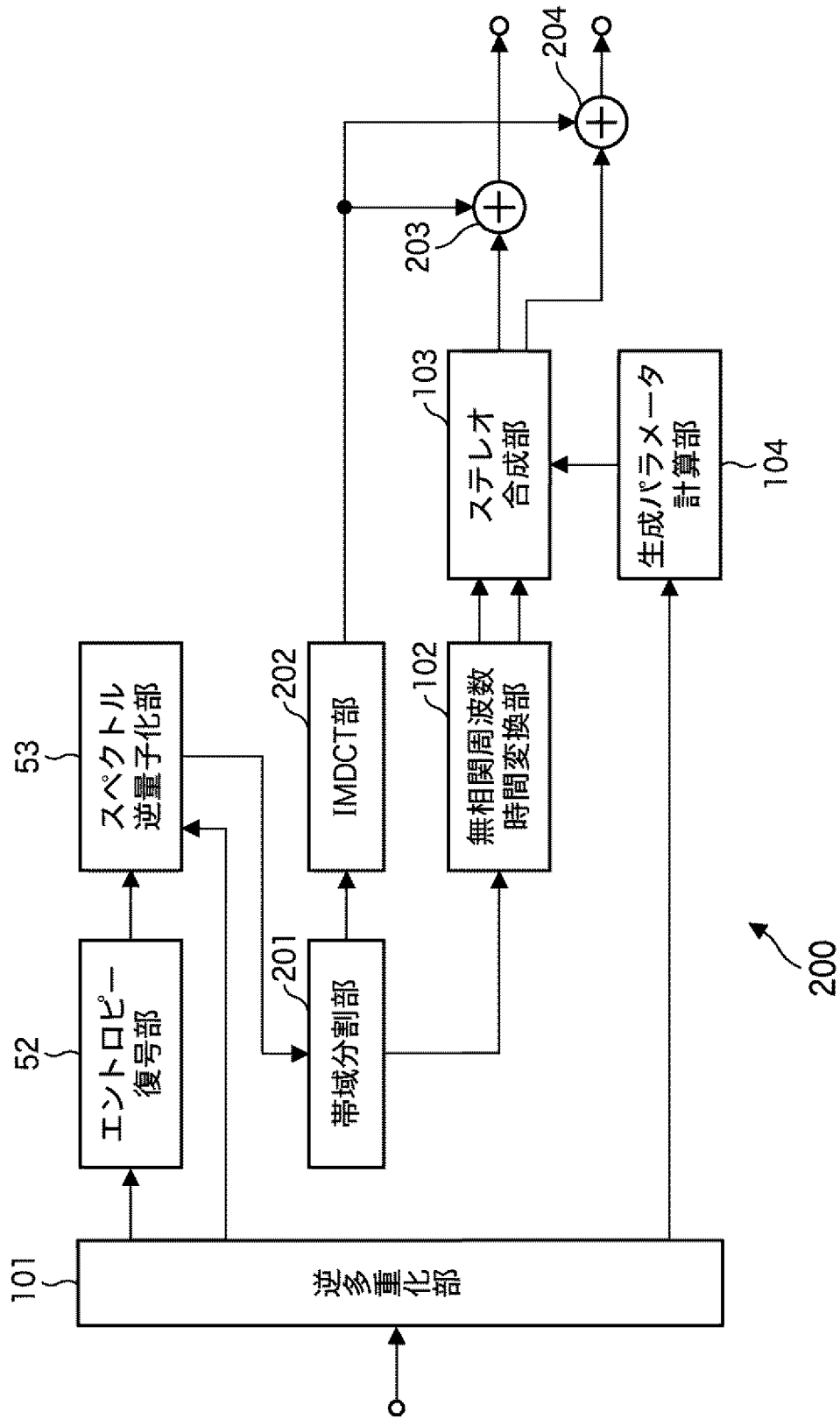


[図14]
図 14



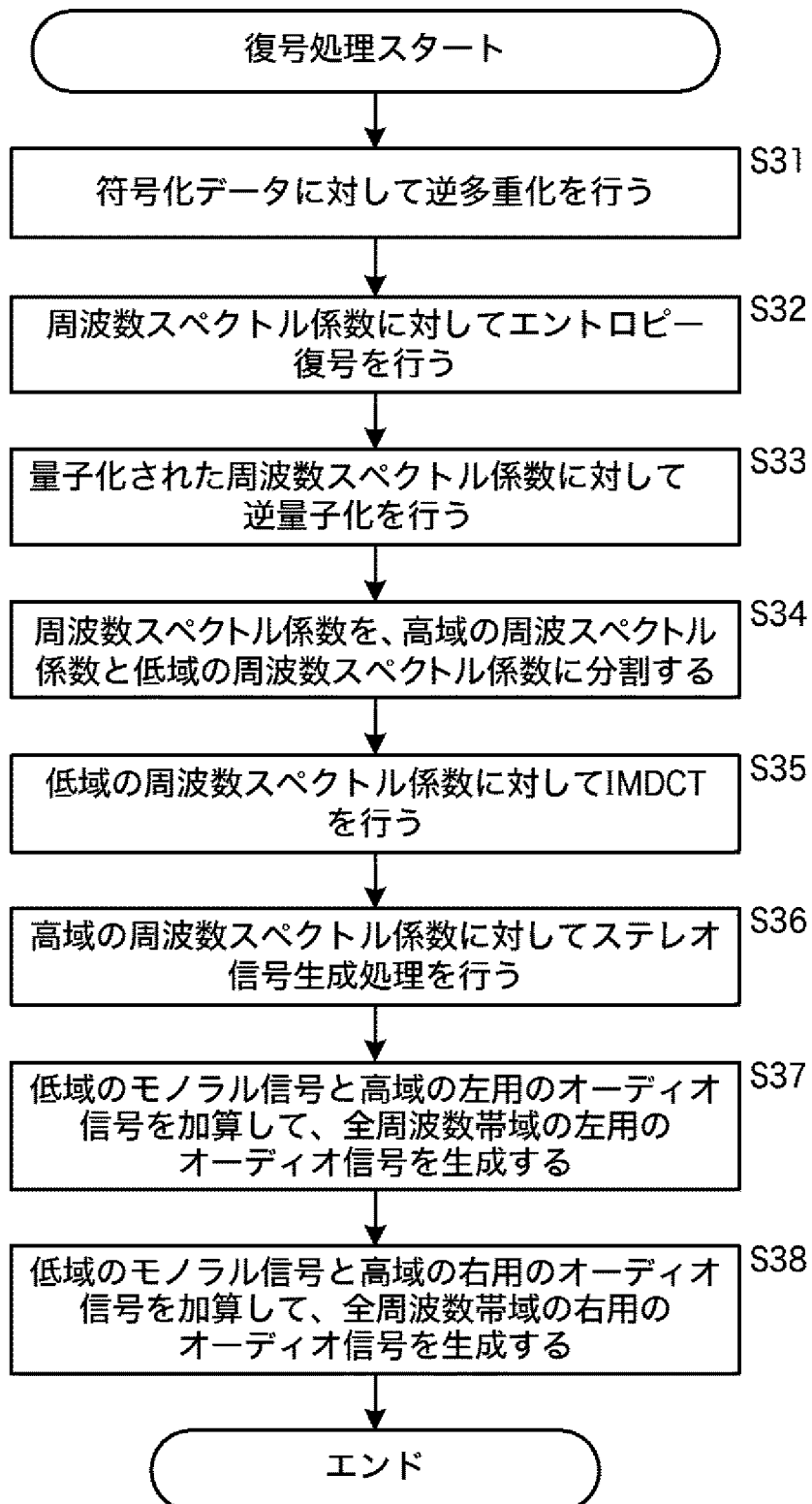
[図15]

図 15

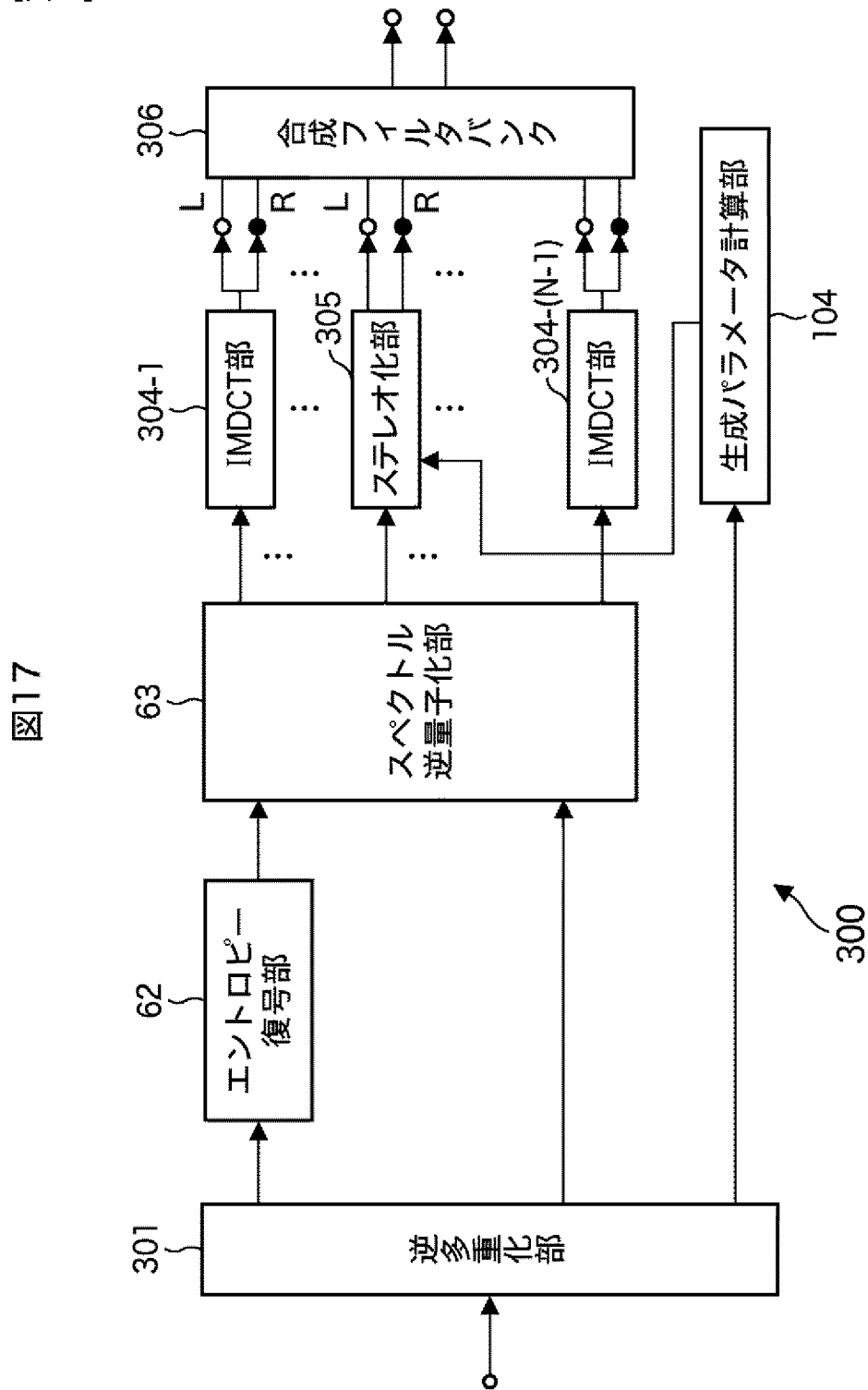


[図16]

図 16

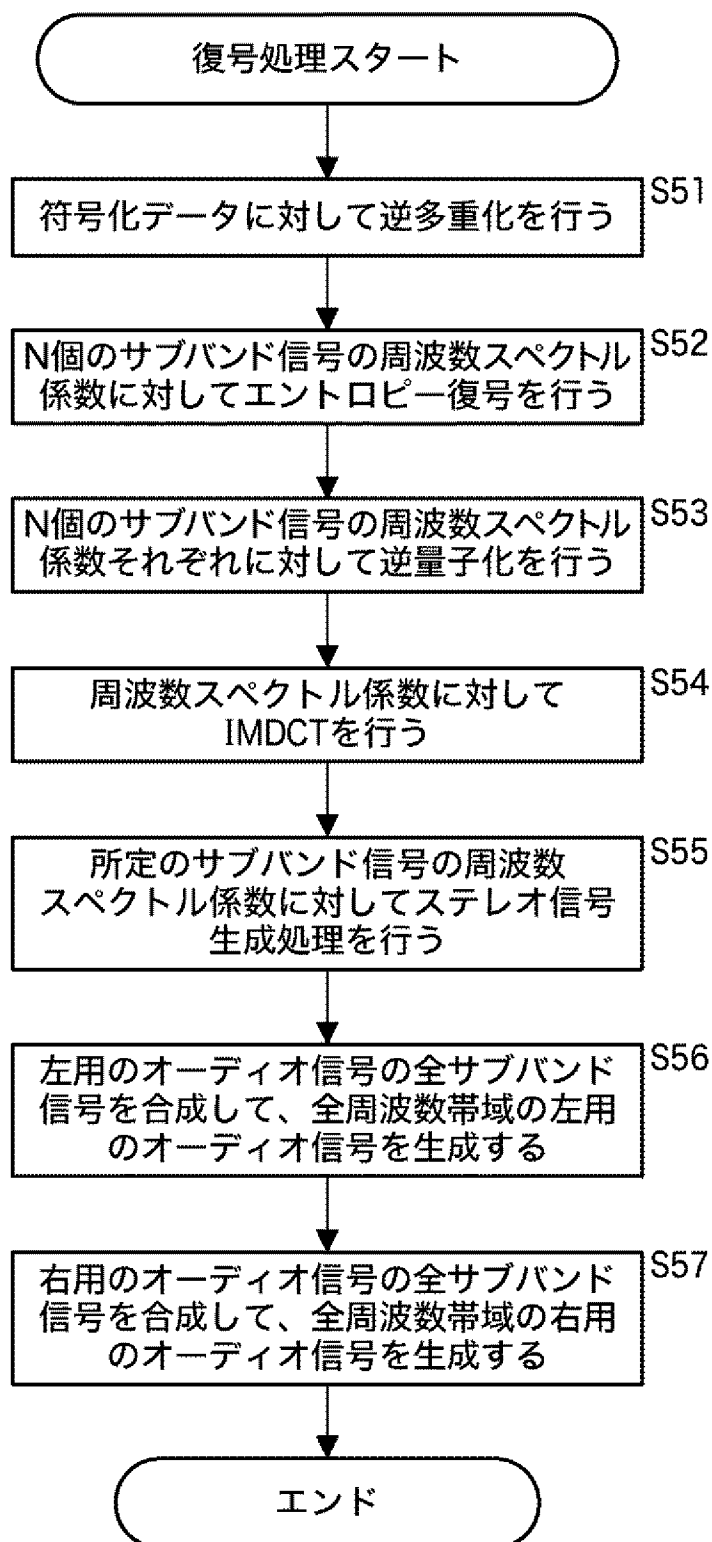


[図17]



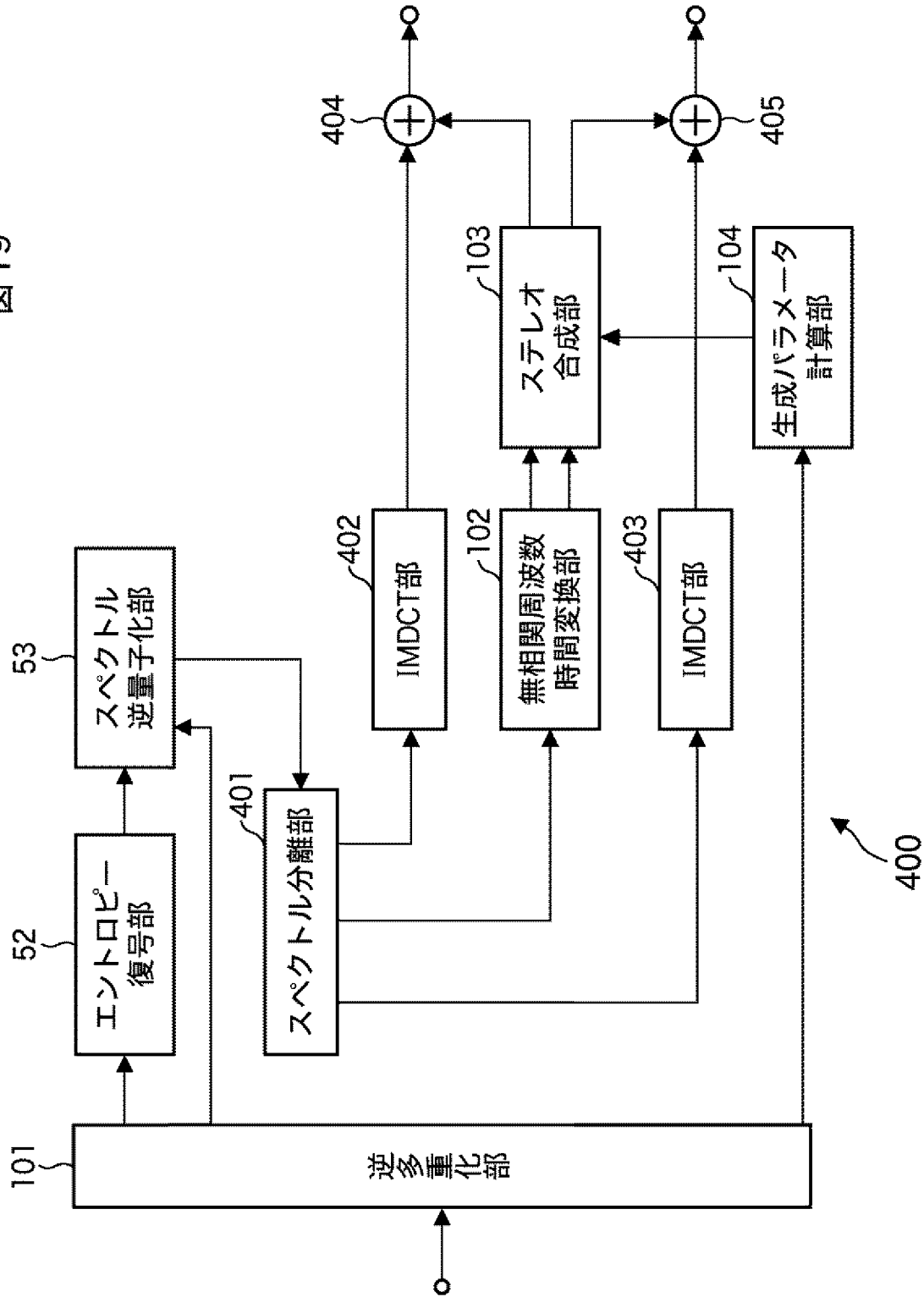
[図18]

図 18

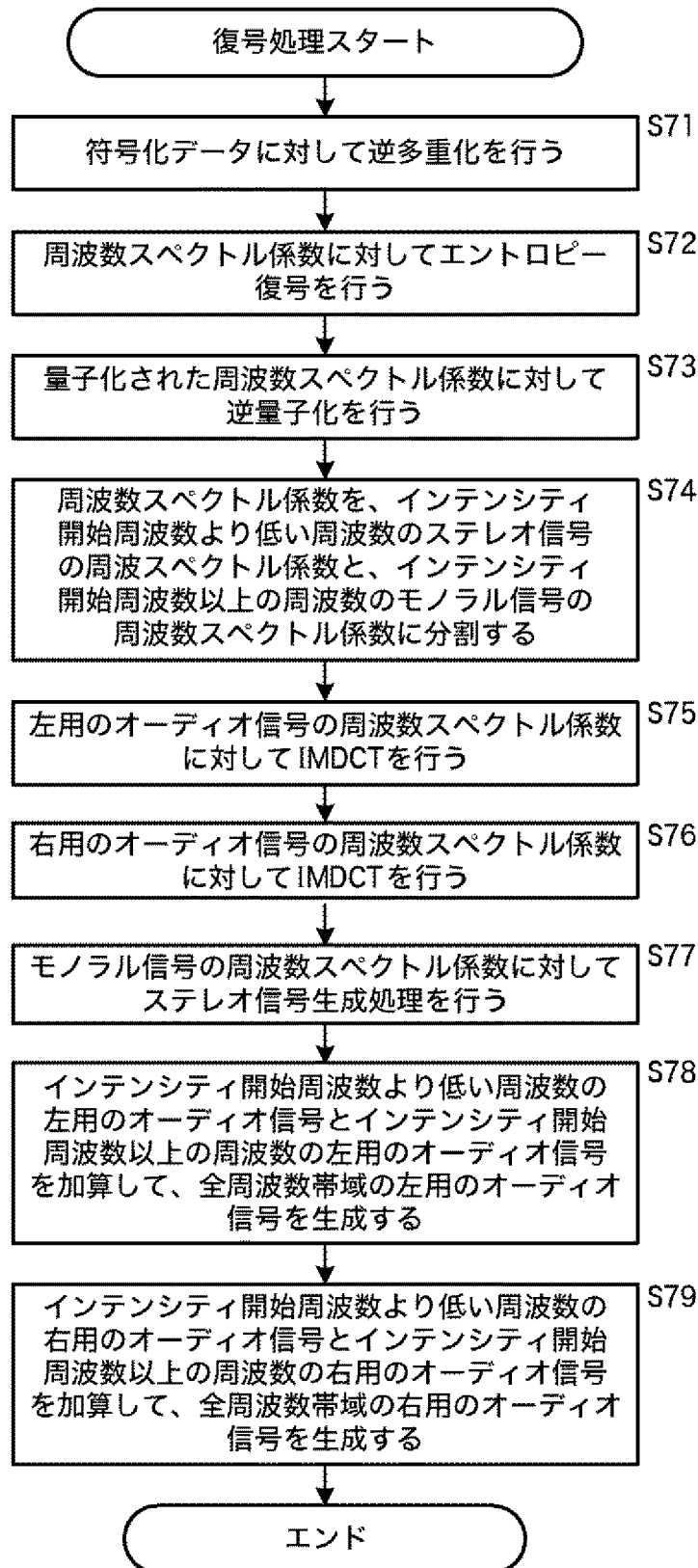


[図19]

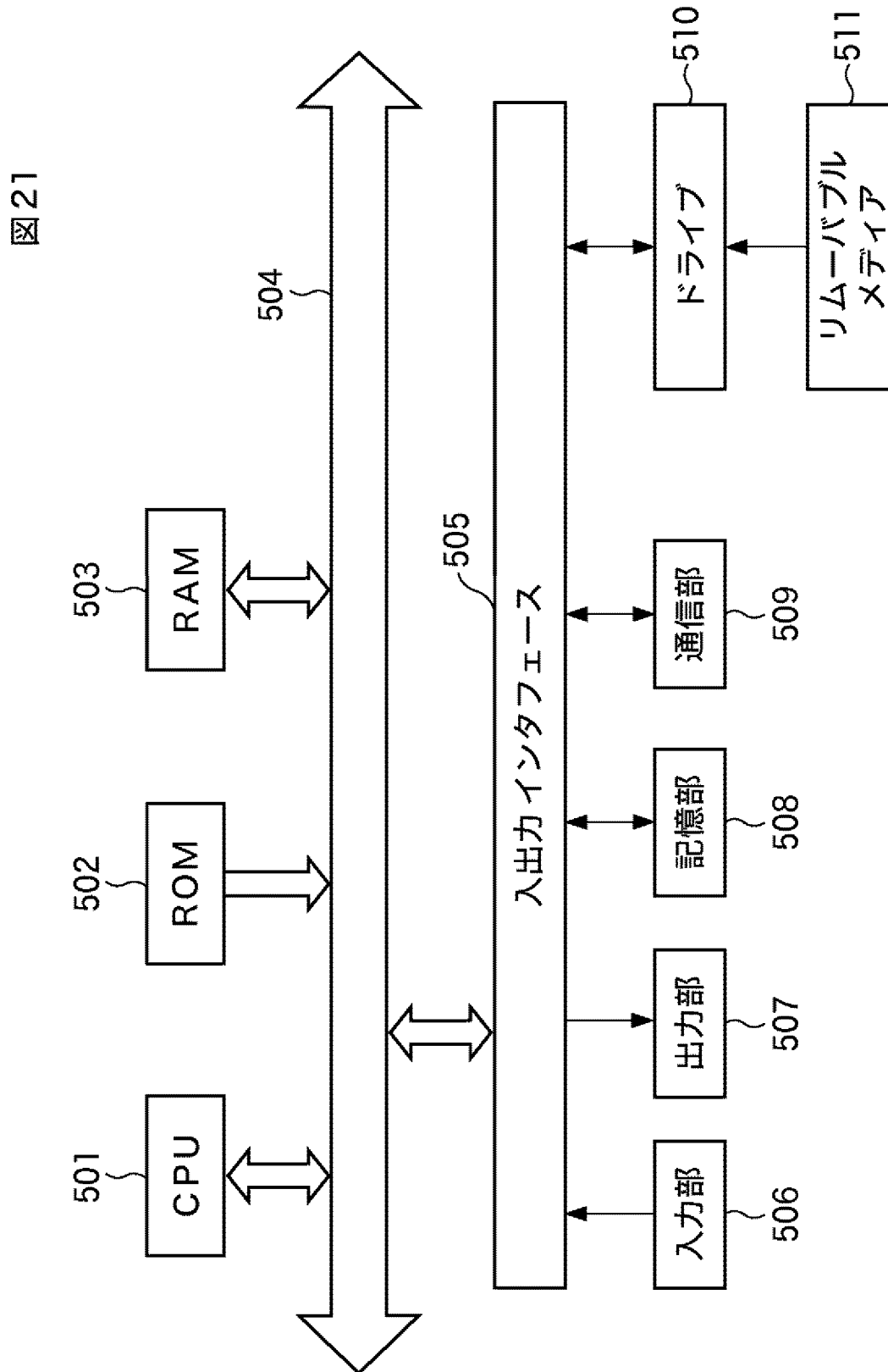
図 19



[図20]
図 20



[図21]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2011/055293

A. CLASSIFICATION OF SUBJECT MATTER

G10L19/00 (2006.01) i, H04S3/00 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L19/00, H04S3/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2011
Kokai Jitsuyo Shinan Koho	1971-2011	Toroku Jitsuyo Shinan Koho	1994-2011

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2007/029412 A1 (Matsushita Electric Industrial Co., Ltd.), 15 March 2007 (15.03.2007), entire text; fig. 1 to 19 & EP 1921605 A1	1-9
A	WO 2007/010785 A1 (Matsushita Electric Industrial Co., Ltd.), 25 January 2007 (25.01.2007), entire text; fig. 1 to 14 & EP 1906706 A1	1-9
A	JP 2006-325162 A (Matsushita Electric Industrial Co., Ltd.), 30 November 2006 (30.11.2006), entire text; fig. 1 to 13 (Family: none)	1-9



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search
05 April, 2011 (05.04.11)

Date of mailing of the international search report
19 April, 2011 (19.04.11)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. G10L19/00(2006.01)i, H04S3/00(2006.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（I P C））

Int.Cl. G10L19/00, H04S3/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 1 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 1 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 1 年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	WO 2007/029412 A1 (松下電器産業株式会社) 2007.03.15, 全文 第 1 図-第 19 図, & EP 1921605 A1	1-9
A	WO 2007/010785 A1 (松下電器産業株式会社) 2007.01.25, 全文 第 1 図-第 14 図 & EP 1906706 A1	1-9
A	JP 2006-325162 A (松下電器産業株式会社) 2006.11.30, 全文 第 1 図-第 13 図 (ファミリーなし)	1-9

C 欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の 1 以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

0 5 . 0 4 . 2 0 1 1

国際調査報告の発送日

1 9 . 0 4 . 2 0 1 1

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号 1 0 0 - 8 9 1 5

東京都千代田区霞が関三丁目 4 番 3 号

特許庁審査官（権限のある職員）

前田 祐希

5 Z

2 9 4 6

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 9 1